Local adaptation under demographic and genetic fluctuations

Neelofer Banglawala

Doctor of Philosophy
The University of Edinburgh
2010

Abstract

Evolution frequently plays out over ecological timescales. Local adaptation under the joint action of evolutionary and ecological processes frequently leads to novel outcomes, as is evidenced by the theoretical work on adaptation at species' borders. However, to date this body of work does not have a theory for the effect of stochastic processes on local adaptation.

The primary goal of this thesis is to show that demographic and genetic fluctuations can significantly impact upon local adaptation. In addition, the effect of polygenic evolution is also analysed. Specifically, three types of models are considered.

First a deterministic mainland-island, subject to hard directional selection, maladaptive gene flow and density regulation is solved for two different trait architectures: an explicit multilocus trait and a quantitative trait. The maladaptive and adaptive steady states can be bistable. This depends on the underlying architecture of the trait, as well as locus number and ploidy. Source-sink structure can emerge, accompanied by a novel, upper critical threshold above which maladaptation occurs. The most favourable condition for local adaptation occurs for few loci and low migration.

Second, a stochastic version of the mainland-island model is analysed as a diffusion process. This is the central premise of the thesis and is explored by examining properties of the stationary distributions of both trait architectures, and the first-passage properties of the single locus case. It is found that across a range of migration rates that depend on locus number and migrant polymorphism, local adaptation may be reversed or escape from maladaptation becomes possible at varying transition rates. The diffusion model is compared to a similar discrete model. The continuous model is in good qualitative agreement with the discrete model.

Third, the stochastic model is generalised to the infinite island model, which evolves deterministically. Under deterministic dynamics a range of equilibria are possible, depending on whether habitat size varies or is fixed. Multilocus dynamics restrict the conditions for polymorphism. Stochastic dynamics can have potentially detrimental consequences for the persistence of the island population when drift is strong.

The relevance of the stochastic model to border populations is discussed. Although the diffusion process imposes severe constraints on the permissible parameter ranges, it is still able to provide a good qualitative understanding of the impact demographic and genetic fluctuations have on local adaptation.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Neelofer Banglawala*)

Acknowledgements

I am grateful to my supervisor Nick Barton, for his guidance and never-ending store of patience. My heartfelt thanks to Jon, who was there at the beginning. I am grateful to Prof. Ackland, Jamie, Davide, Julien, Bhavin, Mike and Satya for discussions and great theory talks. Wilson and Alexander, your enthusiasm and wisdom never failed to lift my spirits. For the indefinite loan of books and sheet music, I must thank Martin and Richard. Thanks Malcolm, for being a great head of postgraduate studies. Jane, you most definitely are postgraduate secretary extraordinaire.

My work would not have been possible without funding from EPSRC. Access to the powerful Eddie cluster made running lengthy simulations a doddle. Lawrence and Rupert, I hope some of your collective geekiness has rubbed off on me, I owe you for the TeX support.

Inside and outwith the realms of academic life in Edinburgh I have had the good fortune of meeting many great personalities. A few have left an indelible impression. Thank you Arno, Kate, Stephen, Matt, Sarah and Freya for copious amounts of tea, Bob Ross evenings, Manjula dinners, Twin Peaks, that trip to Kinross, mochis, tapioca bubbles, spontaneous music-making, an introduction to poles and a general overdose of fun. Thanks Alex, for the tough training and the many laughs; you are an inspiration. To all those I cannot name specifically, it was a pleasure: fellow PhDers, both past and present, Ashworth/IST folk, bballers, fballers and of course the musical storm.

To the one who shone light on fractals and to the one who wrote of ardent wire, thank you for an immeasurable quantity of wisdom, love and inspiration. To the one who sang from the heart, all this and more, lucky me. Thank you.

To my family, thank you for all your love and unwavering support. You are the best. To my dear parents, this is for you.

Contents

Glossary of Terms

additive genetic variance The total variance due to the additive effect of each gene contributing to trait. (Note that the term 'additive' has been used to mean different quantities in the literature. In reference to a particular gene (locus), we will take it to mean the contribution of that gene to the total additive genetic variance.

asymmetric loci Loci that have different allelic effects on the trait and/or experience different migrant pressure.

demographic stochasticity The random variation in population size caused by the (independent) random birth and death of individuals.

density-dependent population regulation The population increases or decreases by an amount that depends on its current size.

diploid Carries two copies of each chromosome.

fitness The number of offspring an organism produces after one generation e.g. it is the value of the fitness function.

gene flow The transfer of genes between different interbreeding populations.

genetic drift The random change in genotypic (or allele) frequency caused by random variation in organism reproduction.

genetic polymorphism See polymorphism.

genotype The set of alleles carried by an individual.

haploid Carries one copy of each chromosome.

haplotype A particular combination of alleles found in a haploid, i.e. a haploid genotype.

heterogeneous environment Environmental conditions and/or selective pressures vary (discretely or continuously) across the range of the environment.

heterozygote An organism with different alleles at a genetic locus.

homozygote An organism with identical alleles at a genetic locus.

linkage disequilibrium Nonrandom associations between alleles at two or more genetic loci.

linkage equilibrium Alleles at different loci are combined at random. Genotypic frequencies are then the product of the frequencies of constituent alleles.

locus A location on the genome, often taken to be synonymous with a gene.

maladaptation Inadequate adaptation or the failure of adaptation altogether.

maladaptive That which prevents adaptation i.e. causes maladaptation to occur.

Malthusian fitness The fitness of an organism that reproduces randomly throughout time (continuous time).

migration The movement of organisms between different populations.

migration load The decrease in average fitness relative to the maximum fitness due to migrants carrying disadvantageous alleles.

natural selection The process which causes genotypes that confer a higher fitness to organisms to increase in frequency.

niche The set of conditions favourable to survival of the species.

random drift The collective term for genetic drift and demographic stochasticity, that is, the random variation in population number and allele frequency due to the random variation in organism reproduction.

recombination The breakdown of combinations of genes and the consequent generation of new combinations.

symmetric loci Loci that are identical; they contribute the same additive amount to trait and experience the same migrant pressure.

Wrightian fitness The fitness of an organism that reproduces in non-overlapping generations (discrete time).

Symbols

N_u	Population size (or population density, since the areal extent of an island is kept constant).
p_j	Frequency of the '+' allele at locus j on the island.
p	Frequency of '+' when only one locus contributes to trait or when loci are symmetric.
n	Number of loci.
s_u	Selection coefficient.
r_0	Intrinsic population growth rate per capita, i.e. the growth rate per capita in the absence of selection and migration.
K	The intrinsic carrying capacity or maximum (equilibrium) population size in the absence of selection and migration.
$\overline{r}_u$	Average Malthusian fitness or average growth rate per capita (continuous time).
$\overline{W}_u$	Average Wrightian fitness (discrete time).
$\hat{p}_j$	Frequency of the '+' allele at locus j in the mainland population.
$\hat{p}$	Migrant allele frequency of the '+' allele when only one locus contributes to trait or when loci are symmetric.
M_u	Number of migrants per generation from the mainland.
m	Rate of gene flow ($= M_u/N_u = M/N$).
$\hat{r}$	Average Malthusian fitness or average growth rate of mainland migrants.
η	Ploidy parameter that has the value 1 for diploids and 1/2 for haploids.
ζ	Scaling parameter ($= r_0/K$), also called the large system size parameter.
N	Scaled population size ($= N_u/K$).
s	Scaled selection coefficient ($= s_u/r_0$) for one locus or symmetric loci.
$\overline{r}$	Scaled average Malthusian fitness or scaled average growth rate ($= \overline{r}_u/r_0$).

M Scaled number of migrants per generation from the mainland ($= M_u/\zeta$), also called the migration constant.

z' Trait.

z Trait mean.

v_z Trait variance.

$\hat{z}$ Trait mean on the mainland.

Z Scaled trait mean ($= \frac{z}{\sqrt{v_z}}$).

$\hat{Z}$ Scaled migrant trait mean ($= \frac{\hat{z}}{\sqrt{v_z}}$).

α_j Allelic contribution of locus j to the trait.

S Scaled selection coefficient in trait-demographic dynamics ($= s\sqrt{v_z}$).

α Allelic contribution of a locus when only one locus contributes to trait or when loci are symmetric.

β Selection gradient; the slope of the graph of fitness against trait.

M_l Lower critical migration rate defined by migration-selection balance.

M_h Upper critical migration rate defined by growth-selection balance.

$\triangle M$ Bistable range or width determined by the deterministic mainland-island dynamics ($= M_h - M_l$).

J_{ij} Jacobian matrix of the deterministic dynamics ($= \partial_{x_j}\dot{x_i}$).

v_j Additive genetic variance at locus j ($= p_j(1-p_j)/2\eta$).

$\mathbf{A}$ (A_i) Deterministic field vector of a diffusion process.

$\mathbf{B}$ (B_{ij}) Diffusion matrix of a diffusion process.

U Potential U of the deterministic dynamics.

$\Psi_n(\mathbf{x})$ Stationary probability distribution (for n loci).

Φ_n Unnormalised stationary probability distribution (for n loci).

$\phi_{N,n}$ Marginal distribution of N (for n loci).

$\phi_{p,n}$ Marginal distribution of p (for n loci).

$\tilde{\Psi}_j$ Approximation of the stationary probability distribution (subscript j labels the type of local peak approximation used, e.g. Gaussian).

Z_0 Normalisation constant.

$P(k)$ Probability of being near peak (or zone) k.

M_{max} Migration rate at which the stochastic population is equally likely to be well-adapted or not.

$\triangle_\zeta M$ Bimodal range i.e. range of migration rates for which the stationary probability distribution is bimodel ($= M_{max} \pm \delta M$).

$T_{met}(N_0)$ Mean extinction time under neutrality or for a single selected locus given initial population size N_0.

$\pi_s(p_0,\, N_0)$ Fixation probability for one locus, given initial allele frequency p_0 and initial population size N_0.

$\pi_0(p_0)$ Fixation probability for one locus under neutrality, given initial allele frequency p_0.

N'_e Estimated effective population size from the fixation probability.

N_e Effective population size.

Γ_{i0} Transition rate from zone i to the saddle 0.

τ_i Mean first passage time out of zone i.

ρ_i Fractional size or frequency of habitat i, where $i \in \{1,\, 2\}$.

N_i Scaled population size of demes in habitat i.

p_{ik} Frequency of '+' allele at locus k in habitat i.

s_i Selective coefficient of '+' allele in habitat i.

γ Relative selective disadvantage of the '+' allele in the foreign habitat (i.e. $= s_2/s_1$).

$\overline{r}_i$ Average Malthusian fitness of habitat i.

$\hat{x}$ e.g. the average population size of the migrant pool $\hat{N}$.

Ψ_i Stationary probability distribution of demes in habitat i (for one locus).

ρ_H Upper critical fractional habitat size, above which the '+' allele spreads throughout the metapopulation.

ρ_L Lower critical fractional habitat size, below which '−' allele spreads throughout the metapopulation.

Chapter 1

Introduction

Natural populations are dynamic entities; they grow, adapt, invade new habitats, and ultimately go extinct. Historically, evolution was thought to mostly take place over timescales much greater than those of ecological processes. Consequently the study of adaptation has largely ignored population dynamics. Likewise, ecologists have been content to ignore that organisms are genetically different. The central theme of this thesis is the theoretical study of how population fluctuations affect local adaptation (adaptation in subdivided populations). This effort is justified by a burgeoning body of evidence that points to evolution occurring over the short timescale of several hundred years, known as contemporary evolution. Furthermore, theoretical studies of adaptation at species' borders model ecological and evolutionary processes over commensurate timescales. However, a theoretical treatment of fluctuating populations is absent from this modest but growing literature. Finally, population fluctuations can cause extinction. Maintaining adaptation, genetic variation and preventing the extinction of endangered populations are some of the goals of conservation and conservation biology.

In this thesis I will show that demographic stochasticity, together with genetic fluctuations, can have a significant impact on local adaptation, in particular when genetic multilocus evolution is coupled to population dynamics. To do this I will explore three different but related theoretical models of local adaptation in heterogeneous environments.

In Chapter 2, I review literature relevant to my thesis. I focus on theoretical eco-evolutionary studies of species' ranges, whilst also providing a more general background, covering island models and population fluctuations in theoretical population genetics.

In Chapter 3 I examine the first of the three models. This is a simple deterministic mainland-island model, in which population size and allele frequencies jointly evolve under gene flow and directional selection. I choose directional selection because it is ubiquitous in nature, and is the simplest form of selection to analyse. This model is the foundation for the other two models. I solve the model for one locus, Sec. 3.3, and multiple loci, Sec. 3.4, under degrees of maladaptive gene flow and assesses the conditions most favourable for local adaptation and persistence of the population at large sizes. To further assess the impact of multilocus evolution on local adaptation, in Sec. 3.5 I contrast the genetic model with a quantitative trait model, for which the genetic details are unknown. Multilocus evolution does indeed give rise to markedly different outcomes from the quantitative trait model. I will show that the number of genes that contribute to the trait and the contribution of each gene to trait are key factors in determining favourable conditions for local adaptation; they can drive the emergence of source-sink dynamics. Furthermore, the nontrivial interaction between demography, density regulation, migration and selection gives rise to a second critical threshold rate, in contrast to classical models that assume constant population sizes, for which there is only one critical threshold.

The second model is the main focus of my thesis and is examined in Chapter 4. I explore a stochastic version of the deterministic mainland-island model that includes demographic and genetic fluctuations. I use the theory of diffusion processes to explore this model and the corresponding quantitative trait model from the previous chapter. By exploring the various stationary and transient properties of the single locus genetic model in Sec. 4.3, I will show that demographic and genetic fluctuations give rise to a very different picture of local adaptation to one without demographic fluctuations. The results are extended to the multilocus case in Sec. 4.4. Again, multilocus evolution has a nontrivial effect on local adaptation. In Sec. 4.5 I explore the stochastic quantitative trait model and find that, analogous to the deterministic model, the absence of the underlying genetic dynamics alters the outcomes for local adaptation. I will also consider how well the diffusion model approximates more realistic discrete models. It may be that the diffusion process requires unrealistic biological parameter values.

The third and final model is an extension of the stochastic mainland-island model to a population with an infinite number of islands in Chapter 5. For this model, I focus on the single locus and multilocus cases. By examining both the deterministic version of the infinite island model in Sec. 5.3, I will again show that multilocus evolution dynamics can significantly impact upon

the favourable conditions for local adaptation. Finally, I will consider the infinite island model under stochastic dynamics and show that the inclusion of demographic fluctuations can cause extinction, in contrast to previous models.

Lastly, in Chapter 6 I discuss the overall approach used to model demographic and genetic stochasticity, in addition to discussing the specific results from each chapter. Possible future work is summarised here.

Chapter 2

Local adaptation in heterogeneous environments

In this chapter I first introduce the basic model that I explore and develop in chapters 3, 4 and 5 of the thesis. I then review the background literature relevant to my work.

2.1 Introduction

Adaptation is driven by natural selection and is the process by which organisms become "better-suited" to their environment. Local adaptation occurs in an heterogeneous environment, where there are selective differences between habitats (known as divergent selection); it is the process by which different populations of a single species, that may be connected by migration and gene flow, become "better-suited" to their *local* environment (i.e. habitat) (Kawecki & Ebert 2004). For example, the snail *Capaea nemoralis* has different shell banding-patterns and colour ratios appropriate for camouflage within its particular habitat (Cain & Sheppard 1950). Evolution under (divergent) selection, however, is rarely a solo act, and crucially depends on other evolutionary and ecological processes, which may hinder or promote local adaptation (Kawecki & Ebert 2004). The failure of local adaptation is important in species' range studies. Small, fragmented and fluctuating populations at the edge of a species' range may experience high migration from central populations that have adapted to a different environment (Kawecki 2008). This migration from central populations may prevent local adaptation in peripheral populations and therefore may prevent the species' from expanding into habitable territory. My work extends the current theoretical

understanding of the failure of adaptation in the species' range literature. I do this by studying three related models of local adaptation in a (discrete) heterogeneous environment, as described next.

2.1.1 The basic model

The first model of local adaptation is considered in Chapter 3. It is based on the mainland-island model introduced by Haldane (1931), which is reviewed in §2.2.1. The basic model is depicted in Fig. 2.1 and is as follows.

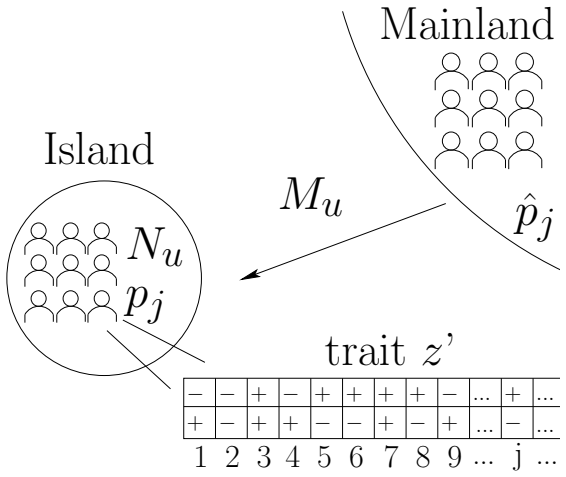


Figure 2.1 *The basic mainland-island model that is developed in the next three chapters of the thesis.*

Consider an environment that consists of an island and a mainland. At any given point in time, the island population (of constant areal extent) consists of N_u individuals. Each individual has a particular value of a trait, z', for example body weight. The trait is made up of k loci (genes), where each locus has one of two alleles, either a '+' or '−' allele. Selection acts on the population such that the locally favoured allele, say the '+' allele, confers a higher fitness to the individual carrying it. Therefore an individual with '+' alleles at all loci will have maximum fitness. The allele frequency of the advantageous '+' allele is p_j at locus j. The population size changes according to the average growth rate of the population. In continuous time models, the average growth rate is also known as the average

Malthusian fitness of the population (Crow & Kimura 1970). Therefore, the higher the average fitness, the greater the average growth rate and the larger the population size will be. However, the population cannot grow indefinitely, hence we introduce density-dependent population regulation: population growth is limited by a factor that is proportional to its current size. Therefore the larger the population, the greater the penalty on its growth rate. This may occur in natural populations, for example, due to overcrowding or a lack of resources. This is known as density-dependent population regulation (Andrewartha & Birch 1954, Murray 1989). Migrants arrive on the island at a rate M_u from a mainland population that is fixed in size and genetic composition. Because migrants are foreign to the local environment, they introduce disadvantageous genes to the island population. Therefore, migrants mostly carry the '$-$' allele (which may confer a high fitness to migrants on the mainland, but not on the island). The frequency of the '$+$' allele at locus j amongst immigrants is $\hat{p}_j$. Throughout the thesis, all migrant quantities are denoted with a hat, unless otherwise stated. The greater the number of disadvantageous alleles carried by migrants (the lower frequency $\hat{p}_j$ at a given locus), the more each migrant hinders adaptation.

We can follow the dynamics of individual allele frequencies explicitly, or we can follow the average trait dynamics, ignoring the dynamics of the underlying genetics. We are interested in both scenarios and consider how N_u and p_j, or N_u and the average trait, vary in time under the processes of selection, growth, density-dependent population regulation, migration (the movement of individuals) and gene flow (the movement of genes). (There is a distinction between migration and gene flow because even if individuals migrate between populations, gene flow only occurs if migrants mate with native individuals so that migrant genes enter the native gene pool.) Ultimately, we wish to determine the conditions (for example, high migration or low maladaptive gene flow) under which the locally favoured '$+$' allele, (or similarly the average trait), can exist on the island at high frequency. In other words, we wish to determine the conditions which favour local adaptation.

In Chapter 3 the basic mainland-island model described above is explored, assuming that the population size and allele frequencies (or average trait) vary deterministically over time. This is appropriate in the limit of infinite populations (i.e. very large populations), in which the stochastic effects of random births and deaths are negligible (Lande et al. 2003). The model reduces to Haldane's mainland-island model if only one locus contributes to trait and the population remains constant. Extensions of Haldane's mainland-island model

which explicitly include deterministic population dynamics, have been explored relatively recently in a subset of the species' range literature. This literature is reviewed in detail in §2.3.1. We build on this work by considering multilocus evolution, directional selection (selection which favours one allele over another). We also contrast the results from modelling the underlying genetics explicitly (i.e. allowing allele frequencies to evolve) with results from following the model with approximate dynamics of the average trait.

The main theme of the thesis is developed in Chapter 4, where we build on the basic mainland-island model by including stochastic effects that result in the population size and allele frequencies (or average trait) fluctuating in time. The stochastic effect on population size is calleddemographic stochasticity (Lande et al. 2003), and that on the gene frequencies is known as genetic drift (Crow & Kimura 1970); the source of both is simply random births and deaths of individuals over time. Since marginal populations are characterised by being small, fragmented and prone to extinction (Kawecki 2008), the effect of stochasticity is crucial to understanding local adaptation. Little work has been done on the effect of stochasticity on local adaptation at the edge of the species' range, see §2.3.1. Outside the species' range literature, Wright quite early on considered fluctuating gene frequencies in his island model but whilst assuming the population size remains constant, see §2.2.2. Wright's assumption is typical of most of the theoretical work on local adaptation carried out in population genetics. Therefore, the work developed in this chapter is relevant more broadly, at the intersection of population genetics and population biology. The assumption of constant population size that is almost exclusively made in population genetics is unrealistic. Similarly, population biology models do not account for the role of genetics in shaping ecology. The stochastic mainland-island model relaxes these assumptions and therefore is a more realistic model of eco-evolutionary interactions than has been considered previously. The role of fluctuating population size in population genetics is briefly considered in §2.4.

Finally, in Chapter 5 the mainland-island is replaced by an infinite number of similar islands. Each island experiences migration from all other demes as if all other demes constitute a mainland population. This seems the most straight forward way to relax the assumption of a fixed mainland (since all demes vary in size and genetic composition), whilst keeping the analytics tractable. Furthermore, migration is no longer unidirectional and is the same between all islands (i.e. it is uniform and symmetric; the distance between islands does not affect the rate of migration). We again only consider two alleles at each locus.

In some demes the '+' allele will be locally favourable, and in others the '−' allele will be favoured. In the case of two alleles, the population of demes (metapopulation), reduces to two distinct habitats, each favouring a different allele. We consider both deterministic and stochastic local dynamics. The island model and two habitat model has a long and extensive legacy. Levene (1953) and others considered the infinite island model and its two habitat variant under deterministic dynamics, see §2.2.1. Even a non-constant population size was taken into account, but only to determine the conditions that favoured the maintenance of both alleles in the metapopulation i.e. the maintenance of genetic polymorphism; the reciprocal effect of adaptation on the population dynamics was ignored. Local adaptation and growth under the joint dynamics of population size and average trait in two habitats under deterministic dynamics was considered more recently; the literature is reviewed in §2.3.1. As for stochastic dynamics, aside from Wright's island model work, to the best of our knowledge, no theoretical work has yet been done to incorporate demographic stochasticity and genetic drift.

2.1.2 Overview of background literature

My work draws from two main themes within the expansive theoretical literature on local adaptation: broadly, the preservation of genetic polymorphism or genetic variation in subdivided populations (Felsenstein 1976, Hedrick et al. 1976, Hedrick 1986), and specifically, adaptation at the edge of the species' range (Pease et al. 1989, Kirkpatrick & Barton 1997, Holt & Gomulkiewicz 1997, Holt et al. 2003, Bridle & Vines 2006, Polechová et al. 2009, Bridle et al. 2010). An important distinction between these focal themes is one of timescales: most studies on local adaptation treat population sizes as fixed, which implicitly assumes that population dynamics equilibrate on a faster timescale than evolutionary processes. Most of the early work on genetic polymorphism and adaptation in subdivided populations makes this assumption. Though this may be justified in certain biological contexts, there is now abundant evidence that evolution can and does occur over ecological timescales (Hendry & Kinnison 1999, Kinnison & Hendry 2001, Saccheri & Hanski 2006, Carroll et al. 2007), for example, the rapid evolution of beak sizes in Darwin's finches on the Galapagos islands (Grant & Grant 2006). Ecological and evolutionary processes can create complex feedback interactions that must be studied in consort in light of the reality of rapid evolution.

A structured population is a subdivided population that has the potential

to experience migration and interbreeding between its subdivisions. The 'subdivisions' have various synonyms: subpopulations, local populations, regions, habitats, islands or demes. If selection on differential fitness does not affect the population density, selection is referred to as 'soft' (Wallace 1975). By 'stable', I simply mean structured populations that do not experience extinctions or recolonizations and that undergo 'soft' selection. Soft selection is selection that does not affect the local population size and so the local population is assumed to remain constant (unless otherwise stated). I concentrate on discrete habitats, in particular island models, and only briefly consider similar work on continuous habitats.

The rest of the chapter is organised as follows. In §2.2, I present an overview of local adaptation in subdivided population when the population size is assumed constant, that is, a 'static' structured population. I focus on the work that was done to determine conditions that maintain genetic polymorphism and genetic variation in the mainland-island setting. These studies characterise the basic interaction between natural selection, gene flow (the movement of genes) and genetic drift (random fluctuation in gene frequencies due to finite sampling), the results of which carry over to more complex models and therefore form the foundation of my work.

I next present a more narrowly focussed review of 'dynamic' subdivided populations in §2.3. By 'dynamic' I mean there may be repeated extinctions and recolonizations, and selection is 'hard' i.e. selection can affect changes in population number (e.g. we would expect a population with high average fitness to have a greater average growth rate than a less fit population). These studies form the specific background to my work. This excludes the extensive literature on the effect of extinctions and recolonizations on genetic variation and fixation probabilities in subdivided populations. The relevant studies are largely drawn from the body of work on adaptation at species' borders, and models of two habitats in which selection is hard (i.e. it directly affects population size). Again, I focus on discrete habitats, and only briefly review the important studies on continuous habitats of species' ranges. This literature is relevant to the work of Chapter 3 and Chapter 5. Importantly, I review what work has been done on demographic and genetic fluctuations; these studies are central to the main theme of my work, developed in Chapter 4.

As mentioned earlier, my work also has relevance beyond species' borders and adaptation in marginal habitats. The drive toward biological realism in evolutionary studies is reflected in the increasing ecological content of models on

local adaptation and the growing evidence of contemporary evolution (Thompson 1998). In §2.4, I briefly review how variation in population size has been treated in population genetics, focussing on what work has been done on random population fluctuations to date. Aside from a small body of work that studies the effect of random population fluctuations on genetic variation and fixation probabilities, little work has been done in population genetics to study models with demographic and genetic drift.

Finally I briefly discuss the implications of my work for conservation biology in §2.5. Note that I do not consider density-dependent selection, which is when genotypic fitnesses depend on population density. This is an important theme in evolutionary ecology and came about from one of the earliest attempts at integrating population genetics and population biology. For more details see MacArthur (1962), MacArthur & Wilson (1967), Roughgarden (1971), Charlesworth (1971), Anderson (1971), Levin (1978), Roughgarden (1979), Slatkin & Maynard Smith (1979), Roughgarden (1979), Christiansen (2004) for more detail. Lastly, this review is meant as an overview only. A more detailed review of the key studies relevant to a particular chapter is given at the start of that chapter.

2.2 Local adaptation in stable structured populations

In stable structured populations, the interaction between natural selection, gene flow and genetic drift, determines the degree to which local adaptation is possible.

Here I briefly discuss the role of gene flow in evolution, before moving on to review genetic polymorphism as a result of natural selection and gene flow in deterministic island models. There is evidence for genetic polymorphism at one locus, as in the well-known case of industrial melanism in cryptic moths (Kettlewell 1973), and most of the literature reviewed deals with one locus only. However, genetic polymorphism in several genes has also been identified, the most well-known example being plants growing on mined soil that have a greater heavy metal tolerance than plants growing on uncontaminated soil closeby (MacNair 1987). Far fewer studies deal with polygenic variation under soft selection, which I also review.

I move on to consider genetic drift and discuss the maintenance of genetic variation in Wright's infinite island model subject to gene flow, genetic drift and selection, reviewing other infinite island models. To conclude the section on

stable population structure, I briefly consider spatially explicit models of clinal variation.

2.2.1 Gene flow and natural selection

Gene flow occurs when migrant alleles that are foreign to the recipient population become part of the local gene pool through mating (Ehrlich & Raven 1969, Lenormand 2002). Gene flow can occur via passive (e.g. the movement of seeds by wind) and active (e.g. migratory birds) dispersal between demes, or by extinction and colonization of entire populations (Slatkin 1985). The importance of its role in adaptation has been controversial (Ehrlich & Raven 1969, Spieth 1974, Endler 1977, Slatkin 1985, 1987). Gene flow can bring genetic variation into demes or spread superior genes between demes, thus facilitating local adaptation (Wright 1931, 1968-78). It can also create a 'migration load' on recipient populations: this is the relative amount by which the average fitness is less than the maximum fitness due to the influx of deleterious migrant alleles. In the extreme case, the migration load is so high that migrant genes 'swamp' the local population (Haldane 1931, Mayr 1963, Slatkin 1985, 1987, Lenormand 2002). Ultimately, the role of gene flow in adaptation is dictated by population structure (the pattern of migration), the mechanism behind gene flow and its relative strength compared to other evolutionary forces. For example, when migrants disperse randomly and the rate of migration is independent of the distance between populations (spatially independent migration), high gene flow is a homogenizing force that reduces genetic variation between demes (Slatkin 1985). However, under varying levels of directed gene flow with a distance-dependent rate, large genetic differentiation can be maintained, as it is for example between Western and Eastern wild great tits, *Parus major*, on the island of Vlieland (Coltman 2005).

Generally, gene flow can overcome natural selection above a certain threshold rate, thus impeding local adaptation. This is maladaptation, which we define to have occurred (at a given locus) if either the favoured allele is lost or it persists at low frequencies. This is a fundamental feature of many of the mathematical models reviewed in this Chapter. Next, the various island models are reviewed.

Haldane's mainland-island model

The simplest deterministic model of selection (hence local adaptation) in a discrete heterogeneous environment is Haldane's mainland-island model, as mentioned earlier (Haldane 1931); it is a simpler version of the basic model

described in §2.1.1. In Haldane's model however, the population size is fixed. Also, only one locus contributes to trait. As before, the '+' allele is locally favoured on the island and has frequency p there. It has a selective advantage s on the island. The fitness of individuals carrying the favoured allele is $1 + s$, otherwise it is 1. The population receives immigrants at a rate m from the mainland population, which is fixed for the deleterious allele, '−', i.e. $\hat{p} = 0$. Note that in this model, since the population size is fixed, m is actually the rate of gene flow; to keep the population size constant, every generation a fraction m of the population is replaced by migrants. The deterministic rate of change of p is given by

$$\frac{dp}{dt} = sp(1-p) - mp. \tag{2.1}$$

The first term on the right is due to selection: the selection pressure s is multiplied by the additive genetic variance, $p(1-p)$ for haploids (Crow & Kimura 1970). The second term is due to migration (Crow & Kimura 1970). Solving this for $dp/dt = 0$, Haldane found that the locally favoured allele could persist in the population at a frequency $p = 1 - m/s$, provided migration is below a critical rate $m_l = s$. That is, adaptation occurs below this rate, albeit at a suboptimal frequency due to the incoming deleterious alleles (i.e. due to the term m/s, which is proportional to the migration load). If $m > m_l$, maladaptation occurs and the favoured allele is lost, $p = 0$. This is the threshold rate between adaptation and maladaptation mentioned earlier, and is a feature of all models of gene flow and selection.

The island model

The island model, a generalisation of Haldane's model, consists of d similar islands connected by migration between all demes (Levene 1953). In his original conception of the model, Levene assumed that diploids from each island are randomly distributed to all other habitats after selection (Levene 1953) i.e. there is maximal mixing between populations. Using a two habitat model, Maynard Smith (1966) proposed that only a fraction m of each population should emigrate each generation (as in Haldane's mainland-island model). The conditions for polymorphism in Levene's island model have been studied by several authors under increasingly general conditions, for example for haploids (Gliddon & Strobeck 1975) and temporal and spatial variations in fitness and selection (Gillespie 1974, 1975). The two habitat model has been studied extensively (Moran 1962, Maynard Smith 1970, Deakin 1966, 1968, Bulmer 1972, Pollak 1972,

Christiansen 1974, Pollak 1974), even under general migration patterns (Karlin 1976). Hoekstra et al. (1985) considered two habitat models with different types of selection and found that the conditions for polymorphism at weak and moderate selection improved, if migration is restricted between demes and selection differs in direction between habitats (Maynard Smith 1970). Heterogeneity also increases chances of an allele persisting (Karlin 1976). Intermediate migration rates may favour polymorphism when selection varies in time and space (Gillespie 1974).

Barton & Whitlock (1997) considered an island model in which an allele has selective advantage s in a fraction (or size) ρ of the entire population and a disadvantage $-\gamma s$ in the remainder. The two fractional groups of demes represent two different habitats. This is similar to the infinite island model described in §2.1.1. Barton & Whitlock found that the allele can be established (persist at high frequencies) below a critical migration rate or for habitat sizes greater than a critical fractional size ρ_c. A similar condition in terms of the ratio of selective differences, γ, was first derived for two habitats by Bulmer (1972).

These models examine genetic polymorphism at one locus. There are far fewer studies of polygenic variation in the island model. Crow et al. (1990) studied a two habitat model with two-way migration and several loci contributing to fitness. Barton (1992) considered a similar model in which multiple loci with identical effects on fitness (such loci are often referred to as being "symmetric"). (These models were motivated by questions other than the preservation of genetic variation in the island model.) In both models a critical rate could be found above which one habitat is swamped by migration irrespective of the selection model, whether migration is symmetric or the number of loci and the fitness of migrants. Barton (1992) found that in the "symmetric" model, where recombination rates are high so that associations between loci (linkage disequilibrium) are negligible, if selection is weak the critical migration rate can be much smaller than the selection coefficient and is fact inversely proportional to the number of loci contributing to fitness. This holds irrespective of whether migrants are more fit (Crow et al. 1990) or less fit (Barton 1992) than the recipient population.

Spichtig & Kawecki (2004) also consider polygenic variation in a two habitat model in which a polygenic trait is selected in each habitat in opposite directions, but migration is symmetric and time is discrete. The trait consists of a finite number of loci and the fitness function can be convex, linear or concave. They find that at equilibrium either one or other allele fixes and no genetic variance remains, one locus is polymorphic and the genetic variance is low, or all loci are equally polymorphic and the genetic variance is maximal. In contrast to the

single locus models, if selection favours heterozygote haplotypes, polymorphism will be lost. Genetic polymorphism is not maintained for high dispersal, weak selection or loci with unequal effects. This is because in their model, the trait mean and variance can evolve separately so that even when the trait mean has an intermediate value, the genetic variance can be low. However, if selection is strong and migration is weak, genetic polymorphism is maintained, similar to the single locus models.

Nonrandom associations between loci strengthen the effect of spatially varying selection so that simultaneous genetic variation at linked loci is more likely (Christiansen & Feldman 1975). Barton (1992) found that linkage disequilibrium had little effect on the threshold migration rate; however the time between introduction and establishment of the selected combination of genes is reduced (see the "hypergeometric" model; also Shpak & Kondrashov (1999)).

Several key points can be drawn from these studies. First, limited migration favours polymorphism, because the migration load in each deme is small enough for the locally favoured allele to persist. Second, the balance between migration and selection gives rise to a critical migration below which adaptation occurs. Third, the greater the heterogeneity between habitats (e.g. the greater the difference between selective pressures between habitats), the more favourable the conditions for genetic polymorphism (see Felsenstein 1976, Hedrick et al. 1976, Slatkin 1985, Hedrick 1986 for detailed reviews of these models; see also Hedrick 2006). Genetic polymorphism has been studied for a wide range of ecological and genetic effects, see Hedrick 1986.

2.2.2 Genetic drift and Wright's island model

The results of the last section are appropriate for very large populations. However, finite populations are subject to genetic drift due to finite random sampling (Ewens 1979, Crow & Kimura 1970, Barton et al. 2007). In an isolated population, genetic drift eliminates all genetic variation from a population at a rate proportional to the inverse population size, until one allele is fixed. However, in a subdivided population with low gene flow between demes, genetic drift prolongs the maintenance of genetic variation (Wright 1931, 1932).

Wright considered an infinite island model, where each deme of fixed size N receives a fraction m of immigrants randomly sourced from all other demes per generation. This is basically Haldane's mainland-island model with genetic drift: the infinite population size assumption means that averages across all demes

remain constant so that each deme experiences migration as if from a fixed, mainland source (Wright 1931, 1932, 1968-78). Each deme is a Wright-Fisher population: mating is random, generations are nonoverlapping and offspring are generated by fitness-weighted binomial sampling (Wright 1931, 1932).

In the absence of selection, when m is small and N is large, if p is the allele frequency, Wright's measure for genetic variation between demes, is the widely used F-statistic, $F_{ST} = \frac{var(p)}{\langle p\rangle\langle q\rangle} = \frac{1}{1+4Nm}$ (Wright 1931), where $\langle p\rangle$ and $var(p)$ are the allele frequency average and variance across demes respectively (the angle brackets denote an average across demes; $\langle q\rangle = 1 - \langle p\rangle$). It is a measure of the proportion of genetic variation between subpopulations (subscript 'S' for 'subpopulation') relative to the genetic variation across the total population (subscript 'T' for 'total'; see Whitlock & McCauley (1999) for review of F_{ST}). Two general results have been determined for Wright's island model. First, there is little genetic differentiation between islands if $4Nm \gg 1$ (assuming symmetric gene flow between demes). Second, at equilibrium the average homozygosity (the total frequency of homozygotes) of a random sample of genes in the subdivided population is independent of the migration pattern and depends on the population size only (Bodmer & Cavalli-Sforza 1968, Crow & Maruyama 1971, Maruyama 1974). See Charlesworth et al. (2003) for a review on the effect of geographic structure on neutral variation.

With selection included, if one locus contributes to trait, the stationary allele frequency distribution $\phi(p)$ can be found by approximating the discrete Wright-Fisher model as a diffusion process (Ewens 1979) in the limit of large population size, when the fluctuations in gene frequencies are small. The use of diffusion processes is discussed in greater detail in Chapter 4, where it is used to describe the coupled dynamics of n loci and population size (see §2.1.1). The basic idea is to formulate a forward Fokker-Planck equation (Gardiner 1985, Van Kampen 2007), which describes the time evolution of the time-dependent allele frequency distribution, from the average dynamics of the allele frequency and the source of noise i.e. genetic drift. The average dynamics of the allele frequency is simply the deterministic equation of its rate of change. This is similar to Eq. 2.1 from Haldane's mainland-island model,

$$\frac{dp}{dt} = p(1-p)\frac{d}{dp}\Big(\ln \overline{W}\Big) + m(\hat{p} - p), \tag{2.2}$$

where $\hat{p}$ is the allele frequency of migrants at the j^{th} locus and $\overline{W}$ is the Wrightian fitness (i.e. defined as the fitness of individuals in discrete time; in this case the fitness defined in the Wright-Fisher model). The term $\partial_p \ln \overline{W}$ gives the selection

pressure on p (Barton et al. 2007, see Ch. 15). We denote the deterministic contribution by $A(p) = dp/dt$. The contribution of genetic drift can be found by considering the Wright-Fisher model: it is the variance of the allele frequency due to binomial sampling with replacement, given the allele frequency in the parental generation (Ewens 1979) and is given by $B(p) = p(1-p)/2N$ for diploids. (Note this is the additive genetic variance for diploids divided by the population size, cf. §2.2.1. The factor 2 appears in the denominator because diploids carry two copies of each gene and therefore the gene pool is twice as large as the number of diploid individuals, whereas haploids only carry one copy of each gene). The one-dimensional Fokker-Planck equation of the time-dependent allele frequency distribution (conditional on the initial frequency p_0 at time t_0), $\phi(p,\, t)$

$$\partial_t \phi(p,\, t) = -\partial_p\Big(A(p)\phi(p,\, t)\Big) + \frac{1}{2}\partial_p^2\Big(B(p)\phi(p,\, t)\Big). \tag{2.3}$$

The stationary allele frequency distribution $\phi(p)$ can be found by setting Eq. 2.3 to zero and solving. At equilibrium, we assume the probability current $J(p) = A(p)\phi(p) - \frac{1}{2}\partial_p\Big(B(p)\phi(p)\Big)$ is zero (Gardiner 1985). This gives a generic form for the stationary distribution, $\phi(p) = \exp\Big(2\int^p A(p')/B(p')\,\mathrm{d}p'\Big)$, provided $B(p) > 0$ for all p (excluding the boundaries $p = 0,\, 1$ in this case). Wright found (Wright 1986) the allele frequency distribution (for diploids) to be

$$\phi(\mathbf{p})_W \propto \overline{W}^{2N} p^{4Nm\hat{p}} q^{4Nm\hat{q}}, \tag{2.4}$$

where $\hat{p} = 1 - \hat{q}$. Unfortunately, it is difficult to generalise this distribution to a finite number of demes and there are no general results (Felsenstein 1976, Hedrick 1986).

Barton & Whitlock (1997) extend Wright's mainland-island model to two habitats subject to directional selection and genetic drift. This is similar to the deterministic island model discussed in the previous section: an allele is favoured in a fraction of demes and is deleterious in the remainder. They found that genetic polymorphism will always be maintained provided $m < s$, irrespective of how low Nm is. However, with genetic drift, if $Nm \ll 1$, the range of critical fractional sizes for which polymorphism can exist also decreases.

From these studies we find that genetic drift promotes genetic variation if migration is restricted in Wright's island model. However, strong genetic drift restricts polymorphism if one allele is favoured in a fraction of demes.

The stepping stone model

The spatial range of migration is limited in nature and populations are likely to exchange migrants only with nearby subpopulations. A more realistic model is the stepping-stone model (Kimura & Weiss 1964), in which migrants are exchanged between adjacent populations distributed linearly or in a planar array. Because migration is spatially limited, high gene flow does not prevent extensive genetic differentiation in the one-dimensional stepping stone model, as it does in the island model (see Maruyama 1971, 1972, Felsenstein 1976 for more details).

2.2.3 Continuously varying selection

Real populations sampled over space for one or more loci show clinal variation (i.e. a continuous gradient) in the gene frequency (Hedrick et al. 1976). For example, allele frequencies for the genes controlling enzyme alcohol dehydrogenase in *Drosophila melanogaster* show clinal variation that increases (decreases) with distances from the equator (Oakeshott et al. 1982).

Haldane (1958) first considered divergent selection in a geographic continuum (but see (Fisher 1937) for uniform selection): an allele is selectively favoured in one half of the range and is deleterious in the other half. Gene flow and selection cause the gene frequency to assume a clinal profile between the boundary of the two regions. Fisher and others considered models in which selection varies smoothly along the continuum (Fisher 1950, Endler 1973). The size of the region in which the allele is favoured is critical to its establishment. For example, if an allele has selective advantage s, then it can be established if this region is greater than a characteristic lengthscale $l = \sigma/\sqrt{2s}$ (Slatkin 1973), where σ is the dispersal range (variance of migration displacement). Similar expressions are found for models of linearly varying selection (e.g. Nagylaki 1975), and quantitative polygenic traits (e.g. Slatkin 1978). Furthermore persistence of the allele at high frequencies becomes less likely outside this region, see e.g. Barton (1987). More details on clinal selection can be found elsewhere, (Felsenstein 1976, Endler 1977, Barton 1999). Similar to the island models, the key result is that, generally, if a critical migration rate is exceeded, the allele cannot be maintained. In the case of spatially distributed populations, this means that local pockets of adaptation may collapse (Felsenstein 1976).

We have discussed soft selection models of heterogeneous environments and the effect of selection, gene flow and genetic drift on genetic variation and the

conditions necessary for genetic polymorphism. In general, migration should be limited between demes, but populations should be large, otherwise genetic drift may restrict polymorphism. Linkage and a highly heterogeneous environment also promote genetic polymorphism.

We next look at how local adaptation under hard selection and variations in population size differs from the soft selection paradigm.

2.3 Local adaptation in dynamic structured populations

Structured populations dynamically vary due to local processes and exogenous factors such as environmental disasters that cause extinction. Hard selection in a heterogeneous environment can have apparently contradictory effects on local adaptation. For example, a hard selection version of Levene's island model was examined by Christiansen (1975), who found that the conditions for genetic polymorphism are restricted compared to the soft selection results. On the other hand, Walsh (1984) showed that the conditions for the invasion of one allele when rare could be less stringent under hard selection, where the number of emigrants from a deme is a function of the deme's fitness. However, these models pay scant attention to the ecological consequences of local adaptation.

In species' range studies, ecological and evolutionary processes are modelled on the same timescale. This goes beyond simply modelling hard selection because the ecological consequences of adaptation are taken into account. We next briefly introduce the species' range problem and the demographic effect of migration, before discussing the different population structures and the key results of local adaptation in dynamic populations structures.

2.3.1 Adaptation at species' borders

Species' ranges exhibit a rich variety of geometries, from interconnected patches, to large, contiguous areas with gaps and multiple peaks of abundance and depletion (Andrewartha & Birch 1954, Brown & Lomolino 1998, MacArthur & Wilson 1967, Hanski & Gilpin 1997, Hanski 1999). Typically, a species distribution is characterised by an abundant interior region with a fragmented, less dense boundary consisting of smaller, fluctuating populations (Lawton & Woodroffe 1991, Lawton 1993, Brown 1995, Thomas & Kunin 1999, Fortin et al. 2005).

The failure of a species to expand into available territory (Brown & Lomolino 1998) may be due to the the failure of border populations to adapt to local conditions (Kawecki & Ebert 2004, Bridle & Vines 2006). There are several possible explanations for this. Border populations, being small and therefore subject to inbreeding, genetic drift, environmental and demographic stochasticity (Bradshaw 1991, Willi et al. 2006), may lack enough genetic variation to respond to selection (Blows & Hoffman 2005). However, even if genetic variation is abundant, adaptation at the edge of the species' range can still fail (Kawecki 2008). Asymmetric gene flow from central populations that have adapted to different conditions can swamp peripheral populations, preventing local adaptation (Mayr 1963, Antonovics 1976, Kawecki & Holt 2002, Polechová et al. 2009, Kawecki & Ebert 2004). Empirical evidence for maladaptation due to gene flow has been found in local populations of patchy habitats (Camin & Ehrlich 1958, McNeilly 1968, Endler 1977, Macnair 1981, Dhondt et al. 1990, Dias 1996) and in particular source-sink habitats, for example in blue tit and mosquito fish populations (Stearns & Sage 1980, Blondel et al. 1992).

Under eco-evolutionary dynamics, gene flow can still swamp an island population. However, once demography is coupled to evolution through migration, the mainland-island can turn into source-sink structures (Harrison & Taylor 1997, Pulliam 1988, 2000). Source populations are typically large and well-adapted. They lie within the species' niche (Hutchinson 1958, Wiens & Graham 2005, Holt 1996*a*, Holt & Gomulkiewicz 2002, Holt 2009), which is the set of conditions favourable to the species' survival. Simply, it is the habitat within which the population has positive growth rate. In contrast, sink populations are poorly adapted with negative growth rates and therefore lie outside the species' niche. They depend on immigration from the source population. Absolute sinks will deterministically go extinct with immigration (Pulliam 2000), whereas pseudo-sinks will persist but at low sizes without migration (Watkinson & Sutherland 1995). Other sink types have been defined: "black-hole" sinks do not send out emigrants, where as "leaky" sinks do (Holt & Gomulkiewicz 1997, Gomulkiewicz et al. 1999).

As we will see, one of the key distinguishing features of the eco-evolutionary models reviewed, in contrast to the soft selection models, is that the rate of gene flow is not constant: it is the ratio of migrants to the recipient population size. Migration couples population dynamics to genetic evolution, creating positive feedback between population size and allele frequency, which leads to results markedly different from the soft selection models discussed in the previous section.

With density-dependent regulation, this can lead to indirect positive feedback between migration and population sizes. For example, "migrational meltdown", when the rate of maladaptive gene flow increases as the population size decreases, further increasing the rate of gene flow, eventually leading to extinction (Ronce & Kirkpatrick 2001).

The demographic effect of immigration has a positive aspect to it. It boosts populations that would otherwise be at risk of extinction. This is known as the "rescue effect" (Brown & Kodric-Brown 1977, Stacey et al. 1997, Hanski 1999). However, the positive role of gene flow does not always accompany its negative role, as for example in the stream populations of three-spine stickleback, *Gasterosteus aculeatus* L., (Moore & Hendry 2009)

The feedback between demography and selection can have a profound effect on both the ecological and evolutionary future of the species. Evolutionarily, novel adaptation may arise in sink populations, being outside the species' niche. This could be important for future adaptability of the species. Ecologically speaking, if sinks can evolve despite maladaptive gene flow, then the species' niche will have expanded. Therefore sink populations are important to niche evolution and niche conservatism (when sinks remain maladaptive) (Kawecki 1995, Holt 1996*a*,*b*).

It is clear that eco-evolutionary dynamics lead to complex interactions. In the remainder of this section I review theoretical work on adaptation at species' borders. Again, I concentrate on discrete habitats, parallelling the review of soft selection models in the previous section. I start with deterministic environments, moving on to briefly consider continuously varying environments. Finally, I will review what work has been carried out on the inclusion of demographic and genetic drift.

One-way migration in source-sink systems

Haldane was the first to identify that positive feedback between gene flow and population density could cause maladaptation in marginal habitats, thus limiting a species' geographic range (Haldane 1956). Several authors have explored this idea by examining the joint, deterministic evolution of population size and allele frequency in source-sink environments in the mainland-island model, as described in §2.1.1 Holt (1983*b*,*a*), Holt & Gomulkiewicz (1997), Gomulkiewicz et al. (1999), Holt & Gomulkiewicz (2002). In these models, a single, diallelic locus undergoes directional selection. The constant mainland population sends out a number of immigrants to the sink per generation; however, unlike the basic model described earlier in §2.1.1, the sink population would go extinct (i.e. the sink is an absolute

or "black-hole" sink (Pulliam 1988, Gomulkiewicz et al. 1999)). Migrant genes are deleterious in the sink population and therefore the environment is maximally heterogeneous.

These models show that adaptation in the sink is possible when the favoured allele is rare, provided its absolute fitness is greater than one (Holt & Gomulkiewicz 1997). Without density-dependent regulation, this criterion is the only one that ensures adaptation; in contrast to the soft selection results, the rate of migration (gene flow) does not matter (Holt & Gomulkiewicz 1997). With density-dependent population regulation, large population sizes depress the average fitness in the sink (Gomulkiewicz et al. 1999). As in soft selection models, low migration rates, below a critical threshold, are favourable for adaptation. In contrast to soft selection models, at intermediate migration rates more than one steady state is stable, which can lead to genetic polymorphism (Gomulkiewicz et al. 1999). Adaptation in the sink can lead to niche evolution, whereby the sink can survive independent of migration from the source (Gomulkiewicz et al. 1999, Holt 1996*a*, Holt et al. 2003). This is an example of niche evolution (Holt 1996*a*, Holt et al. 2003, Wiens & Graham 2005, Holt et al. 2005, Holt 2009). Lastly, unstable population dynamics can severely hamper adaptation (Holt 1983*b*,*a*).

These results carry over to polygenic evolution in source-sink environments (Tufto 2001, Holt, Gomulkiewicz & Barfield 2003). For example, Tufto (2001) considered the joint evolution of population size and the average of a polygenic trait under hard stabilising selection (e.g. fitness is a Gaussian-type function of trait) and density-dependent regulation. The authors also found that multiple stable equilibria exist at intermediate migration rates. Holt et al. (2003) found that without density-dependent regulation, adaptation occurs below a critical value of initial maladaptation between source and sink. This contrasts with the single locus result that the absolute fitness of the favoured allele is the only criterion for spread.

Explicit multilocus evolution in source-sink (or mainland-island) models has yet to be explored (though see Tufto 2001 for simulation results). In Chapter 3, we extend the work discussed in this section by modelling explicit multilocus evolution in the basic model described in §2.1.1, as well as a similar quantitative trait version of the same model.

Two habitat models

Adaptation in two habitat environments under hard selection (selection that causes the population size to vary) has been approached in several different ways. For instance, given a set of environmental and genetic constraints, the optimal phenotypic value is sought (Via & Lande 1985, 1987, Van Tienderen 1991, Kawecki & Stearns 1993) or the optimal evolutionary stable strategy (Brown & Pavlovic 1992) and costs associated with adaptation are analogous to the cost of maladaptive gene flow. Another approach is the assessment of habitat performance using only the total fitness averaged across habitats (Holt & Gaines 1992, Kawecki 1995, Holt 1996*a*,*b*). More recently, the conditions for the success of an invading allele using explicit eco-evolutionary dynamics have been investigated (Kawecki & Holt 2002, Kawecki 2000), as well as the evolution of specialists and generalist under joint evolution of an (polygenic) trait and population density (Ronce & Kirkpatrick 2001, Kawecki & Holt 2002, Filin et al. 2008). Evolutionarily, a specialist is a species that is proficient in very few habitats, and is often associated with monomorphism (fixation of the favoured allele). In contrast, an "ecological" specialist is abundant over a narrow range of habitats. Similarly, an "evolutionary" generalist is well adapted across several habitats. Ecologically, a generalist occupies a wide range of habitats.

Similar to the single locus soft selection models (models with constant population size; cf. §2.2), increasing isolation can favour local adaptation within each habitat (which is known as extreme habitat specialisation, Brown & Pavlovic 1992). With source-sink dynamics however, large migration rates are necessary for adaptation to occur in sink habitats, because evolution is biased towards source habitats (Holt & Gaines 1992, Kawecki 1995, Holt 1996*a*,*b*). This is an example of niche conservatism (Holt & Gomulkiewicz 1997, Wiens & Graham 2005, Holt 2009). Kawecki & Holt (2002) bridge the bap between these two seemingly contradictory results by showing that low migration rates are favourable for invasion of a rare allele if the sink habitat is a pseudo sink and can persist without immigration (Watkinson & Sutherland 1995), otherwise large migration rates are favourable. This is the case when source-sink structure emerges due to habitat quality or simply due to asymmetric migration rates (Kawecki & Holt 2002). Effectively the same result was found by Kawecki (2000), who found that a foreign allele with a small selective advantage can invade a sink if migration is large, but if the source-sink structure can switch (which is similar to the condition of a pseudo sink in Kawecki & Holt (2002)), alleles with large selective advantages can invade the sink at relatively low migration rates. If migration is too large however,

positive feedback between demography and evolution can lead to "migrational meltdown" and eventual extinction of the sink habitat (Ronce & Kirkpatrick 2001). Finally, if density dependence is strong at low densities, a generalist species' is likely to evolve (Filin et al. 2008, Kawecki & Holt 2002, Spichtig & Kawecki 2004).

These models show that multiple stable equilibria can appear due to density-dependent regulation of population numbers, similar to the source-sink models discussed earlier. However, these models suggest that intermediate migrations are unfavourable for local adaptation, in contrast to the models discussed in the previous section.

Once again, explicit multilocus genetics have not been considered in two habitat models with eco-evolutionary dynamics. Furthermore, the infinite island model has also not featured in this literature. We explore multilocus evolution in the two habitat model with variable habitat sizes in Chapter 5; this is done using the deterministic island model described in §2.1.1. We also consider the infinite island model with stochastic local dynamics, which has not been done before. We discuss stochastic models after the next section on spatially varying models.

Continuously varying species' range

Pease et al. (1989) were the first to consider the spatial distribution of a species under maladaptive gene flow and hard selection. They investigated how a species' range tracks a moving environmental gradient along which stabilising selection on a quantitative trait varies linearly. They found that either the range expands uniformly or the population goes extinct. Extensions of this basic model include: static environments (Kirkpatrick & Barton 1997, Garcías-Ramos & Kirkpatrick 1997), invasion (Garcías-Ramos & Rodríguez 2002), competition (Case & Taper 2000), evolving genetic variance (including explicit multilocus evolution, Barton 2001) and two habitats Filin et al. (2008). Environments that vary in both space and time have recently been considered (Polechová et al. 2009). From these studies we learn that ranges can disappear, extend uniformly across space or be limited depending on the steepness of the gradient, the dispersal range and/or the speed with which the environmental optimum moves across space. Just as for discrete habitats, peripheral populations can switch between being sinks and sources (Kirkpatrick & Barton 1997, Polechová et al. 2009).

The effect of demographic and genetic drift

We come to the studies that form the basis of the central theme of the thesis, developed in Chapter 4. The models we have considered so far are deterministic and hence implicitly assume very large (or infinite) population sizes. However, real populations do not evolve deterministically, nor are they infinite in extent. As a result, gene frequencies and population sizes randomly fluctuate both in time and space. Ecologists identify two main sources of stochasticity (May 1974). Demographic stochasticity arises from the random birth and death of individuals (Lande et al. 2003). Since this depends on population size, smaller populations are more prone to experiencing demographic stochasticity (Willi et al. 2006). Environmental stochasticity, on the other hand, will affect a population irrespective of its population size (May 1974, Lande et al. 2003). For example, fluctuations in annual rainfall, or catastrophes that may cause whole populations to go extinct.

Peripheral populations are often small and highly vulnerable to extinction (Brown & Lomolino 1998). However, the impact of demographic stochasticity and genetic drift on local adaptation under joint eco-evolutionary dynamics has largely been ignored until recently.

The effect of genetic drift in a sink habitat was first examined by Gomulkiewicz et al. (1999). They found that the probability of the favoured allele persisting in the population for a long time was highest at intermediate migration rates.

A key work that examines the effect of demographic and genetic drift on local adaptation was carried out by Holt et al. (2003). They investigated local adaptation of a quantitative trait in a sink-source system subject to migration, stabilising selection, density-dependent regulation and stochasticity using individual-based simulations. Their main result was that for a range of intermediate migration rates, the population fluctuates about either a maladapted or well-adapted state for long periods, before rapidly shifting to the alternate state. Crucially, stochasticity provides an opportunity for the sink population to escape maladaptation.

Other studies that take genetic and/or demographic drift into account are models of adaptation along an environmental gradient. For example, Alleaume-Benharira et al. (2006) studied the effect of genetic drift in a stepping stone model along an environmental gradient. They found that local adaptation under gene flow was best possible under intermediate migration rates. This is similar to the result found by Gomulkiewicz et al. (1999). Bridle et al. (2010) examined adaptation along a linear environmental gradient in a quantitative

trait model with stabilising selection, density-dependent regulation and stochastic events through varying offspring numbers, mutation and random dispersal. They concluded that the carrying capacity and the dispersal rate were key factors in determining how well the population adapted to the ecological gradient.

Demographic and genetic drift can have a dramatic effect on local adaptation. Although simulation studies have been carried out to investigate the joint effect of these two sources of randomness, the role of demographic stochasticity and genetic drift needs a theoretical basis upon which future investigations can be built. We provide this theoretical treatment in Chapter 4. In particular, we explore multilocus as well as phenotypic (i.e. trait) evolution.

2.4 Fluctuating population size in population genetics

Evolutionary studies usually make the assumption that demographic processes act over a faster timescale than evolutionary forces and therefore a separation of timescales can be employed (Carroll et al. 2007, Pelletier et al. 2009). In other words, population structure is assumed to have equilibrated and therefore the population size can be treated as constant. A more realistic treatment of population variations is afforded by the use of an 'effective population size', N_e. This is the number of breeding individuals in a population. It can be found, for example, as the size of a randomly mating, idealised population under genetic drift that would give the same variance of allele frequencies as the population under consideration (Crow & Kimura 1970, Ewens 1979). In general, the effective population is smaller than the actual population size, given that random mating is unrealistic assumption, see Whitlock & Barton 1997.

Population fluctuations can cause extinction. A metapopulation is a structured population which experiences frequent local extinctions and recolonizations (Hanski 1999). Fluctuating population structure has a profound effect on the genetic architecture of metapopulations and has been studied extensively in the literature with the use of the effective population size, see for example Slatkin 1977, Maruyama & Kimura 1980, Slatkin 1981, Lande 1984, Wade & McCauley 1988, Whitlock & McCauley 1990, McCauley 1991, Lande 1992, Barton 1993, Barton & Whitlock 1997, Charlesworth et al. 2003, Pannell 2003. This large body of work assumes that local population dynamics act on a faster scale than metapopulation dynamics. For example, Barton & Whitlock (1997) extend the island model to include genetic drift and extinction. For low population sizes

(strong genetic drift) and high extinction rates (weak selection), polymorphism is excluded for a narrow range of habitat frequencies. This has yet to be verified with explicit stochastic demographic dynamics.

Investigation of the effect of (stochastic) population fluctuations on adaptation has mainly focussed on fixation probabilities (see Lande et al. 2009 for an alternative to fixation probabilities). Early work assumed that demographic changes were independent of the underlying genetics (Karlin 1968, Heyde & Seneta 1975, Donelly & Weber 1985). More recently, birth-death processes have been used to determine fixation under demographic stochasticity and density regulation (Parsons & Quince 2007*a*,*b*, Parsons et al. 2008), and environmental stochasticity (Engen et al. 2009). Kaitala et al. (2006) did simulation studies on the spatial genetic differentiation of a single locus in metapopulation with spatially limited migration, genetic drift, demographic stochasticity and density regulation. However, theoretical work that fully explores the evolutionary and ecological consequences of demographic stochasticity has yet to be carried out.

2.5 Summary

We have reviewed island model studies in which the population size remains constant (soft selection models); these form the general background to my work. This classic body of work characterises the basic interaction between gene flow, selection and genetic drift from a static structural point of view. Once demography is properly accounted for by including explicit population dynamics (and not simply modelling hard selection), we see that coupled eco-evolutionary dynamics give rise to nontrivial interactions that are largely specific to the details of the models used, but which have a few universal features such as the critical threshold between adaptation and maladaptation.

The work of my thesis is specifically based on the models reviewed in §2.3. A theoretical treatment of demographic and genetic fluctuations goes beyond extending species' range studies. It contributes to the steady drive towards a more interdisciplinary approach to studying biological processes. Moreover, it may have implications for conservation biology, the discipline that relies heavily on evolutionary theory but also deals with extinction. With the myth of incommensurate timescales thoroughly debunked, contemporary evolution is very relevant to conservation biology (Carroll et al. 2007, Pelletier et al. 2009, Kinnison et al. 2007, Kinnison & Jr. 2007, Stockwell et al. 2003), as is the integration of ecological and evolutionary processes for the management of endangered species

(Hedrick 2004, Hedrick et al. 1996, Lande 1988).

Chapter 3

Deterministic mainland-island model

In this chapter I examine in detail the basic mainland-island model described in §2.1.1, subject to deterministic biological processes that couple population dynamics with evolutionary dynamics. In this model, migrants from a mainland population carry disadvantageous genes (maladaptive gene flow) to an island population every generation. However, migrants also boost the island population size. As mentioned in §2.1.1, population numbers are kept in check by introducing a cost to the growth rate as the population becomes large (density-dependent population regulation). I investigate the conditions that promote (hinder) local adaptation, as well as what conditions give rise to large (small) population sizes. I do this by considering two different approaches to the evolutionary dynamics. In the first approach, the dynamics of the individual gene (single locus) or genes (multilocus) that contribute to trait are coupled to the population dynamics and solved for explicitly. In the second approach, if it is assumed that the genetic dynamics are unknown, the evolutionary dynamics are described by the approximate dynamics of the average trait (phenotype). I show that in the genetic model, in addition to the classic migration threshold due to migration-selection balance, as found by Haldane (see §2.2.1), the coupling between growth and evolution gives rise to a second threshold migration rate for local adaptation. I also show that whether all loci contribute identically to trait (loci are symmetric) or not (loci are asymmetric), can have a strong effect on population numbers and adaptation. These results depend sensitively on the evolutionary dynamics assumed. In the quantitative trait (phenotypic) model, the second migration threshold does not exist, demonstrating the fact that different assumptions made

about the evolutionary dynamics can lead to markedly different behaviour of the model. First, I briefly review the literature my work builds upon and extends.

3.1 Short review of previous models

Gomulkiewicz et al. (1999) analysed a discrete-time, source-sink model subject to density-dependent regulation and hard (directional) selection for a diploid locus (see Holt (1983*b*,*a*), Holt & Gomulkiewicz (1997) for earlier work). The authors examined the conditions for spread of the favoured allele in the sink population when it is initially rare. Several key results came out of their work. First, they found that a necessary criterion for the spread of the local allele is that its *absolute* fitness, or equivalently the absolute fitness of its heterozygote, has to be greater than one. This contrasts with the usual assumption made in population genetics, that only the *relative* genotypic fitnesses matter. Second, if fitness is negatively dependent on population size, there are three equilibrium scenarios. At low migration rates the sink is polymorphic and is larger than the intrinsic carrying capacity. When migration is high, the sink is monomorphic for the deleterious allele at low population numbers. If individuals are diploid, both these states are bistable at intermediate migration rates. Third, if the favoured allele is established in the sink, the sink may evolve such that it can persist without immigration from the source, an example of niche evolution (Holt et al. 2005, see Ch. 2).

Trait-population dynamics under stabilising selection have also been studied (Tufto 2001, Holt et al. 2003, Holt & Gomulkiewicz 2002). Multiple loci contribute additively to the trait and recombination rates between loci are taken to be high (acting at a faster rate than selection). Only the dynamics of the average trait are followed: this requires the restrictive assumption of constant genetic variance (Falconer & MacKay 1996). Nonlinear feedback between adaptation and demography gives rise to multiple equilibria provided the initial conditions are such that gene flow is maladapted. With density-dependent fitness, migration must not exceed a critical rate if the sink (or pseudo-sink, see Tufto 2001) population is to locally adapt.

I extend the single-locus model introduced by Holt & Gomulkiewicz (1997), Gomulkiewicz et al. (1999) to multiple loci, for diploids and haploids in continuous time without imposing sink dynamics on the island. I also investigate a quantitative trait version of the model and contrast the results with the explicit

genetic case. As I will show, both models predict markedly different outcomes for local adaptation: with genetic dynamics, local adaptation and the emergence of source-sink structure is dependent on locus numbers, as well as the relative strengths of migration, selection and growth. Local adaptation may also depend on the initial state of the population and the deviation of allelic effects from symmetry. In contrast, for the quantitative trait model, adaptation is more likely with increasing isolation. However, the average trait and population size can 'blow up' at small migration rates.

The chapter is laid out as follows. In Sec. 3.2, I describe the general model. Starting with the genetic model, I describe the general multilocus dynamics. I then analyse the single locus model in §3.3 for the special case when the local allele is lost on the mainland (§3.3.4). I will go into some detail here, since this case is the 'worst-case' scenario for local adaptation. I will show that a second critical migration rate above which local adaptation is not possible emerges: bistability between an adaptive and maladaptive state and source-sink structure are the hallmarks of positive feedback between selection and population dynamics. Strikingly, this means that adaptation can be different for diploids and haploids. I then consider polymorphic gene flow before moving onto the multilocus model in Sec. 3.4. A similar analysis to the single locus case follows for the symmetric multilocus case (loci are identical) in §3.4.2. I will show that the number of loci has a significant impact on the conditions for bistability and hence local adaptation. To conclude my examination of the genetic model, in Sec. 3.4.3, I consider asymmetric loci. By treating the asymmetric equilibria as perturbed symmetric equilibria, I show that large perturbations in the migrant frequency or allelic effects can affect the conditions favourable for local adaptation by either altering or destroying bistability. In Sec. 3.5, I examine the quantitative trait model and show that in contrast to the genetic model, source-sink structure does not emerge (i.e. no bistable equilibrium). This model is more like soft selection models in which the population size is constant. I end the chapter with a discussion of my findings, placing them in the context of previous work.

3.2 The general model

As described and depicted in §2.1.1, consider an island with N_u individuals (the subscript 'u' is for 'unscaled', anticipating the parameter scaling that is to come). We assume the island resides within a constant environment. Furthermore, the areal extent of the island is kept constant, therefore population number and

population density are exchangeable terms. Individuals are characterised by a trait and this trait is subject to selection. Every generation, the local population receives M_u migrant individuals from a large mainland source. The mainland population is assumed to have reached demographic and genetic equilibrium and therefore does not change over time. Also, there is no backward migration from the island to the mainland. We describe further biological details in the next section, where we present the population dynamics and the evolutionary dynamics of the genetic model.

3.2.1 Population dynamics and the genetic model

We assume the trait consists of a number of finite diallelic loci with '+', '−' alleles (see Fig. 2.1 for a depiction of the trait). Each locus contributes a small amount to the trait, the value of which depends on the allelic state of the locus (Bulmer 1980). The sum of these contributions gives the total value of the trait. In the genetic model, the evolutionary dynamics are the dynamics of the individual genes. Let the '+' ('−') allele be selectively advantageous (disadvantageous) and its frequency at locus j be p_j ($q_j = 1 - p_j$). We assume n loci contribute to trait, where $n \geq 1$, as depicted in Fig. 2.1. To follow the genetic dynamics of the trait we would need to track the evolution of each locus and all the associations between pairs of loci, three loci, four loci and so on. For more than a few loci this quickly becomes intractable. However, we can simplify the dynamics by neglecting the nonrandom associations between loci (linkage disequilibrium), if we assume that combinations of genes get broken down at a rate faster then selection (recombination is faster than selection) (Falconer & MacKay 1996). Loci can then be treated as evolving independently of each other in linkage equilibrium, and we can write down an equation for the deterministic rate of change for the j^{th} allele frequency.

The dynamical equations for population size N_u and allele frequency p_j are,

$$\frac{dN_u}{dt_u} = N_u(\overline{r}_u - \frac{r_0 N_u}{K}) + M_u, \tag{3.1a}$$

$$\frac{dp_j}{dt_u} = p_j(1 - p_j)\partial_{p_j}\overline{r}_u + \frac{M_u}{N_u}(\hat{p_j} - p_j) \tag{3.1b}$$

where $\overline{r}_u$ is the average (Malthusian) fitness or average growth rate, $\hat{p}_j$ is the allele frequency of the j^{th} migrant locus, r_0 and K are respectively the intrinsic growth rate per capita and intrinsic carrying capacity (or maximum size) of the

population in the absence of selection and migration.

These equations need some qualifying. Let us first look at the population dynamics, Eq. 3.1a. The first term on the right-hand side is the contribution of evolution (selection) to the growth of the population; the greater the average fitness of the population, the higher the average growth rate per capita. This term couples evolutionary dynamics with population dynamics. Since selection affects the population size it is 'hard' (Wallace 1975). Generally, the average fitness function, $\overline{r}_u$, depends on the underlying genetic or phenotypic assumptions of the model and on the type of selection acting on the population. In the genetic model, $\overline{r}_u = \overline{r}_u(\{p_j\})$ is a function of all the allele frequencies and the selection pressure at each locus. Its specific form will be detailed in the relevant sections later in the chapter. Without some form of population regulation, the population would grow exponentially large. Therefore we introduce a density-dependent term, the second term in Eq. 3.1a. The quadratic dependence on N_u ensures that at small sizes, the population can grow almost unrestrainedly. For example, when few individuals are around, there is no competition for resources and the population will grow rapidly. However, at large sizes, resources may becomes scarce and the population will not be able to sustain a high growth rate. This is known as logistic regulation (Murray 1989, Ch. 2). The fraction r_0/K sets the strength of density-dependent regulation (if large, population regulation is severe). In the absence of selection or migration, if $\overline{r}_u = r_0$, then Eq. 3.1a is the classic logistic equation (Balakrishna 1991). Many forms of density-dependent regulation have been proposed and studied in the literature (Henle et al. 2004). Indeed, we need not choose logistic regulation, any form of density-dependent population regulation will do, provided it is an increasing function of population size. However, we opt for logistic regulation because it is the simplest. The final term in Eq. 3.1a is due to migration.

The dynamical equation for p_j is similar to the equation Haldane derived in the original mainland-island model, see 2.1. The first term on the right-hand side is due to selection (Crow & Kimura 1970). The gradient $\partial_{p_j}\overline{r}_u$ gives the selection pressure on locus j (in terms of Haldane's model, this would be the selection coefficient s in Eq. 2.1). We assume that selection is directional, and therefore $\overline{r}_u$ is a linear function of the trait i.e. all n allele frequencies (this is a generalisation of Haldane's model). However, Eq. 3.1b is quite general and can be used for any type of selection (which will dictate the form of $\overline{r}_u$), provided it does not depend on the population size or the allele frequencies. The second term in Eq. 3.1b is due to gene flow (Crow & Kimura 1970). Immigration introduces foreign genes to the

island gene pool. In Haldane's model, the rate of gene flow is m and is fixed. The rate of gene flow in our model is $m = M_u/N_u$, which varies as the population size varies; migration couples selection with population dynamics. Assuming migrants have perfectly adapted to the mainland environment, the degree to which migrant genes are 'foreign' to the demic gene pool is characterised by $\hat{p}_j$: if $\hat{p}_j = 0$, the mainland and island are maximally (genetically) divergent (at locus j). Note that all loci couple to population size through the average growth rate, $\overline{r}_u$; therefore although loci are treated as independent of each other, they are indirectly coupled through the response of N_u to average fitness.

With the multilocus dynamics specified, we now wish to find the conditions under which the favoured allele can remain frequent on the island and the population can attain large sizes. We start by investigating a single locus, moving on to the multilocus case, focussing on two distinct migrant regimes: when gene flow is (i) totally maladaptive, $\hat{p} = 0$, and (ii) when it is polymorphic (partially adaptive), $0 < \hat{p} \ll \frac{1}{2}$. As we will show, the single-locus results qualitatively carry over to the multilocus case, which is expected since loci evolve independently. However, there are important differences due to the indirect coupling of loci to population size.

3.3 Single locus case

In this section we examine the model when fitness differences are caused by a single locus. To find the criteria for spread of the local allele, we will solve the dynamics and examine the the equilibria for totally maladaptive gene flow, $\hat{p} = 0$. In the latter case, we will go into some detail and show that the equilibrium is bistable for diploid individuals only. We will analytically determine the critical migration rates between which bistability occurs. To conclude the single locus section, we consider polymorphic gene flow. We will show that bistability disappears in this case unless the local allele is rare on the mainland, $\hat{p} \sim 0$.

But first, we derive the functional form of $\overline{r}_u$ for a single locus.

3.3.1 Average fitness $\overline{r}_u$

Suppose the '+' allele has selective advantage $\frac{1}{2}s_u$ (s_u is the unscaled selection coefficient), and that it occurs in the island with frequency p. We can derive an expression for the average fitness $\overline{r}_u$ by considering the equivalent discrete

population with nonoverlapping generations and noting that $\overline{W}_u \approx e^{\overline{r}_u} = 1 + \overline{r}_u$ when $\overline{r}_u$ is small (Crow & Kimura 1970), where $\overline{W}_u$ is the average Wrightian fitness in the discrete population. Specifically, in a diploid discrete population, genotypes $++ : +- : --$ have Wrightian fitnesses $(1 + r_0 + s_u) : (1 + r_0) : (1 + r_0 - s_u)$ at Hardy-Weinberg frequencies $p^2 : 2pq : q^2$ (Crow & Kimura 1970), where $q = 1 - p$ is the frequency of the deleterious allele. We choose these fitness values to ensure that in the absence of selection, the average growth rate will still be positive and nonzero $(1 + r_0)$. Similarly, the Wrightian fitnesses for haplotypes $+ : -$ are $(1 + r_0 + \frac{s_u}{2}) : (1 + r_0 - \frac{s_u}{2})$ at frequencies $p : q$. Provided we can ignore terms of order r_0^2, s_u^2 and $r_0 s_u$ (selection is weak), the average Wrightian fitness is

$$\begin{aligned} \overline{W}_u &= 1 + r_0 + s_u \eta (p - q) \\ &\approx e^{r_0 + s_u \eta (p-q)}, \end{aligned} \tag{3.2}$$

where η is the 'ploidy parameter': $\eta = \frac{1}{2}$ for haploids, $\eta = 1$ for diploids. Therefore the average fitness is

$$\overline{r}_u(p) \approx r_0 + s_u \eta (2p - 1), \tag{3.3}$$

using $q = 1 - p$. How sensitive growth is to selection is set by an 'effective selection coefficient' $s\eta$. Note that we have assigned the Wrightian fitnesses to diploids and haploids in such a way that the selective value of homozygotes $++ : --$ is double that of the equivalent haplotypes $+ : -$. This emphasises the additive selective value of the favoured allele. In classical population genetics, it is usual to assign the *same* selective value to the homozygote and haplotype, so that if one allele is dominant, the homozygotes would have the same selective value as the equivalent haplotypes.

3.3.2 Scaling parameters

Through judicious rescaling of the variables, the number of free parameters can be reduced. Dividing the equations in Eq. 3.1 by the scaling parameter $\zeta = r_0 K$ gives

$$\frac{dN}{dt} = N(\overline{r}(p) - N) + M, \tag{3.4a}$$

$$\frac{dp}{dt} = p(1-p)s + \frac{M}{N}(\hat{p} - p) \tag{3.4b}$$

where $N = N_u/K$, $M = M_u/r_0K$, $s = s_u/r_0$, $t = t_u r_0$ and $\overline{r}(p) = 1 + s(2p-1)$ (note $\partial_p \overline{r} = s$).

3.3.3 Solving the dynamics

We wish to solve the (scaled) dynamics and find the equilibrium states. The coupled, nonlinear dynamical equations admit at most four unique equilibria. The stability, degeneracy and biological validity of each steady state is controlled by the parameters M, s, $\hat{p}$ and η. If we assume N has reached equilibrium, its solution, found from Eq. 3.1a by setting $\dot{N} = 0$, can be substituted into Eq. 3.1b, forming a quartic polynomial, $Q_p(p, \hat{p})$, in p only. A similar quartic, $Q_N(N, \hat{p})$, can be formed for N using the equilibrium solution of p from $\dot{p} = 0$. The roots of these two quartics are the four equilibrium values of N and p that form the equilibria $\{N, p\}$. These quartics are,

$$\begin{aligned} Q_N(N, \hat{p}) =& N^4 - 2\eta N^3 + (1 - 2M + 2\eta M - \eta^2 s^2)N^2 \\ &+ 2M(1 - \eta + \eta^2 s - 2\eta^2 \hat{p} s - \eta^2 s^2)N + M^2(1 - 2\eta). \end{aligned} \tag{3.5a}$$

$$\begin{aligned} Q_p(p, \hat{p}) =& 4Ms^2p^4 + 4Ms(1 + 2s - 3s\eta - 2\hat{p}s)p^3 \\ &+ 4M(M - 1 - \hat{p}s - s^2 + \eta s^2 + 3\eta\hat{p}s)p^2 \\ &+ 4M\hat{p}(s - \eta s^2 - 2M)p + 4M^2\hat{p}^2. \end{aligned} \tag{3.5b}$$

The roots of the quartics can be found exactly. However the expressions are impenetrable. Instead, we assess the special case of totally maladaptive gene flow (maximal environmental heterogeneity), before considering partially maladaptive gene flow.

3.3.4 Totally maladaptive gene flow

How does the island population adapt when the local allele is lost from the mainland $\hat{p} = 0$. This is the 'worst case' scenario: if the island can adapt under totally maladaptive gene flow (all foreign genes are locally unfavourable), it should certainly fare better when the local allele is rare. We first determine

simple analytical expressions for the equilibria, moving on to characterise the bistability that arises for intermediate migration rates. The critical migration rates that delimit this range are important for understanding local adaptation.

The dynamical equations in Eq. 3.4 are easily solved for $\hat{p} = 0$ to give the following equilibrium solutions for the allele frequency (dropping the asterisk),

$$p_h = 1 - \frac{M}{N_h s}, \tag{3.6a}$$

$$p_l = 0, \tag{3.6b}$$

and the equilibrium population size,

$$N_h^{\pm} = \frac{\overline{r}_{max} \pm \sqrt{(\overline{r}_{max})^2 + 4M(1-2\eta)}}{2}, \tag{3.7a}$$

$$N_l^{(\pm)} = \frac{\overline{r}_{min} \pm \sqrt{(\overline{r}_{min})^2 + 4M}}{2}, \tag{3.7b}$$

where $\overline{r}_{min} = \overline{r}(0)$ and $\overline{r}_{max} = \overline{r}(1)$. Substituting $N_h^{(\pm)}$ into Eq. 3.6a uniquely determines p_h.

The relative strengths of migration and selection dictate whether the island population is polymorphic, p_h, or monomorphic for the unfavourable allele, p_l, at equilibrium. This is a classic result and goes back to Haldane's work on gene flow and selection for soft selection (Haldane 1931), as discussed in §2.2. When selection is stronger than migration, a balance is struck between the degrading force of maladaptive gene flow and adaptation. The small migration load results in a genetically polymorphic island and the favoured allele is able to reach potentially high frequencies, p_h. We will also refer to the polymorphic state as the 'high' or 'adaptive' state. If instead migration is stronger than selection, the island is overrun by maladapted migrants and it becomes monomorphic and the favourable allele is lost, $p_l = 0$. The monomorphic state will also be referred to as the 'low' or 'maladaptive' states.

The four steady states are: $\{N_h^{(\pm)}, p_h\}$ and $\{N_l^{(\pm)}, p_l\}$. We can immediately discount state $\{N_l^-, p_l\}$, because N_l^- is negative or complex for all parameter values. The stability and biological validity of the remaining three states, $\{N_h^+, p_h\}$, $\{N_h^-, p_h\}$ and $\{N_l^+, p_l\}$, depend on ploidy and the relative strength of selection and growth.

The effect of ploidy

Ploidy (characterised by η) makes a significant difference to the equilibrium behaviour of the population through its effect on the equilibrium population size, see Eq. 3.7a: for every immigrant, diploids experience twice the cost of deleterious migrant alleles as haploids do, since they carry twice as many homologous genes per locus and the cost is the same *per gene* (cf. §3.3). In other words, diploids are more vulnerable to the degrading effect of maladapted gene flow than haploids. This genetic cost translates into a demographic cost through the average fitness $\overline{r}$, specifically the effective selection coefficient, $s\eta$. Consequently, there are only two biologically valid haploid states (recall that for haploids $\eta = \frac{1}{2}$). Therefore the state $\{N_h^-,\, p_h\}$ is unbiological, because since $N_h^- = 0$, p_h^- is indeterminate. Of the remaining two states, $\{N_l^+, p_l\}$ and $\{N_h^+, p_h\}$, only one can be stable for a given set of parameters. Therefore, we have the interesting result that a haploid population is always monostable at equilibrium for a single locus. On the other hand, for diploids, $\eta = 1$ and state $\{N_h^-,\, p_h\}$ is well-defined. Therefore a diploid population can be bistable as well as monostable at equilibrium; the final equilibrium state depends on the initial conditions of the population.

The effect of ploidy on the stability of the equilibrium states, and consequently its effect on local adaptation and demography, depends on the relative strengths of growth and selection. We therefore next take a closer look at the steady states $\{N_h^+, p_h\}$, $\{N_h^-, p_h\}$ and $\{N_l^+, p_l\}$ when growth dominates selection, and vice versa.

Growth dominates selection

When the effective selection strength satisfies $s\eta \leq 1$, selection is weak relative to growth.

The stable equilibrium population size, allele frequency and average fitness for high growth are shown in Fig. 3.1. As the figures show, at a critical migration rate M_l ($M_{l,\eta}$) the equilibria are discontinuous. This is the threshold rate found by Haldane, determined by the balance between selection and gene flow, at which the local allele is lost. For migration rates below M_l, selection dominates migration and the polymorphic state $\{N_h^+, p_h\}$ is stable. That is, the island population adapts and the locally favoured allele persists at high frequencies. As migration increases, the allele frequency (hence average growth rate) decreases monotonically, until the critical threshold M_l, when the migration load becomes so large that the high state allele frequency p_h collapses to zero $Ns = M$.

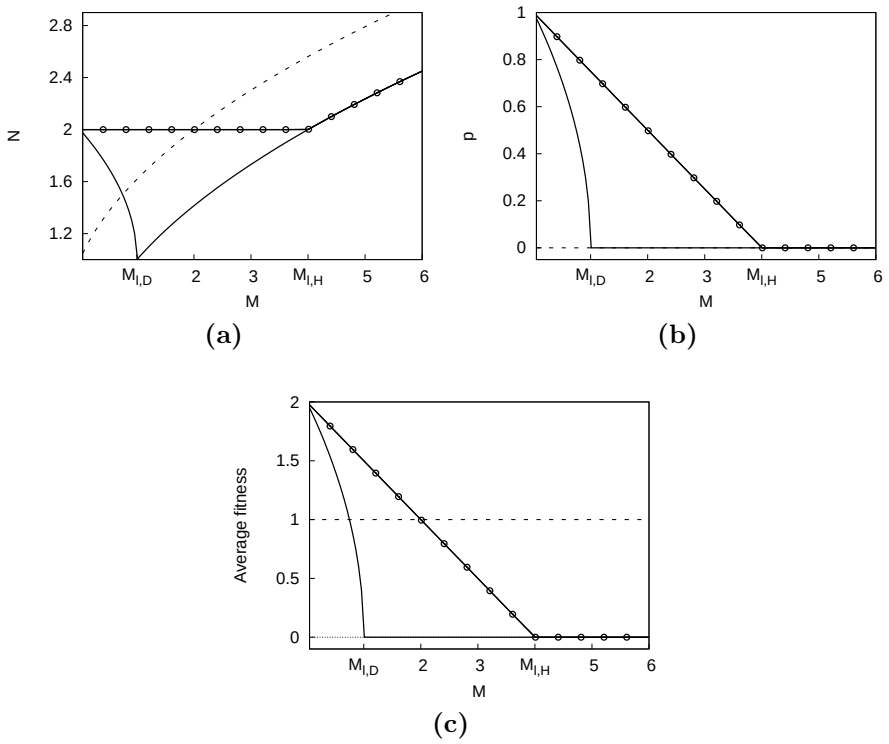


Figure 3.1 *Single locus deterministic equilibria when growth dominates selection: the equilibrium (a) population size, (b) allele frequency and (c) average fitness $\overline{r}$ are shown for diploids (solid lines, $s = 2$) and haploids (with circles, $s = 1$) for $\hat{p} = 0$. The neutral equilibria are also shown for comparison (dashed lines). The critical rates $M_{l,\eta}$ (indicated in the figures) are $M_{l,D}$ for diploids, $M_{l,H} = 4$ for haploids. Critical migration rates M^* (not indicated), below which the favoured allele is more common, are $M^* = \frac{3}{4}$ for diploids and $M^* = 2$ for haploids.*

Above M_l, selection is weaker than migration and the maladaptive, low state $\{N_l^+, p_l\}$ is stable. The local allele is lost as migrants swamp the island, just as in Haldane's mainland-island model (see §2.2.1). Additionally, the average growth rate becomes zero. Setting $N_l = M_l/s$ and rearranging, the critical rate is,

$$M_l = \begin{cases} s & \eta = 1, \\ s(\frac{1}{2}s + 1) & \eta = \frac{1}{2}. \end{cases} \tag{3.8}$$

The critical rate M_l is lower for diploids than it is for haploids. This is because diploids are more susceptible to maladaptation than haploids, as discussed earlier. The ploidy effect is also the reason that, for migration rates below M_l, the haploid population size remains constant despite the increasing migration load and the steady depletion of the favoured allele, see Fig. 3.1a; the increase in selective deaths (decrease in $\overline{r}$) due to gene swamping is balanced by the number of

incoming migrants, see Eq. 3.7a. In a diploid population, incoming migrants cannot compensate for the greater loss in average fitness. Once the island ceases to experience gene flow, ploidy no longer affects demography, as can be seen for migration rates above M_l, when the monomorphic steady state is stable.

An important consequence of strong growth is that despite maladapted gene flow the average growth rate does not fall below zero, see Fig. 3.1c. For the average fitness to fall below zero, the allele frequency must satisfy $p < \frac{1}{2} - \frac{1}{2s\eta}$. Since p cannot be negative, this is satisfied only if selection is stronger than growth, i.e. $s\eta > 1$.

The neutral equilibria are also shown in Fig. 3.1. Whilst the island population is polymorphic, as migration increases the allele frequency falls below $\frac{1}{2}$; neither allele is more common than the other on the island and the average growth rate is the same as the neutral average growth rate, $\overline{r} = 1$. The migration rate at which this happens ($p_h = \frac{1}{2}$) is

$$M^* = \frac{s}{4}(2 + s), \tag{3.9}$$

which is ploidy-independent. Therefore, a high degree of adaptation occurs for low migration rates, $M < M^* < M_l$, when the local allele is more frequent on the island than the deleterious allele and the average growth is higher then the neutral value.

Once migration dominates selection, the island is completely maladapted. Despite this genetic cost, there is no associated demographic cost as such as; the population remains at least as large as the intrinsic carrying capacity, irrespective of ploidy or the strength of migration relative to selection. Above the threshold rate M_l, the population stops growing, but the population size continues to increase due to high migration.

The results of this section follow simple intuition: if growth is a relatively stronger force than selection the population will be saved from low average growth rates despite maladaptation, which occurs once migrant genes dominate the gene pool, for migration rates below M^*. In the next section we consider the opposite case, when selection dominates growth.

Selection dominates growth

If selection is stronger than growth, $s\eta > 1$, the main difference with the previous scenario is that the average growth rate can become negative and hence the local population can become a sink; it collapses to small sizes and requires migration to increases its size. In Haldane's model, the population size was constant,

and therefore it could not become a sink. In our model, migration couples the dynamics of population size and genetic change, creating positive feedback between the two. The more maladaptive the migrant genes (i.e. the smaller $\hat{p}$), the greater the reduction in the population's average fitness. The population size then reduces because adaptation influences the population size through the average fitness. A smaller population size increases the rate of maladaptive gene flow into the deme, which further diminishes the population size, completing the feedback. This positive feedback is a type of "migrational meltdown" (Ronce & Kirkpatrick 2001). Migration also increases the population size, which is known as the "rescue effect" (Hanski 1999, Ch. 4).

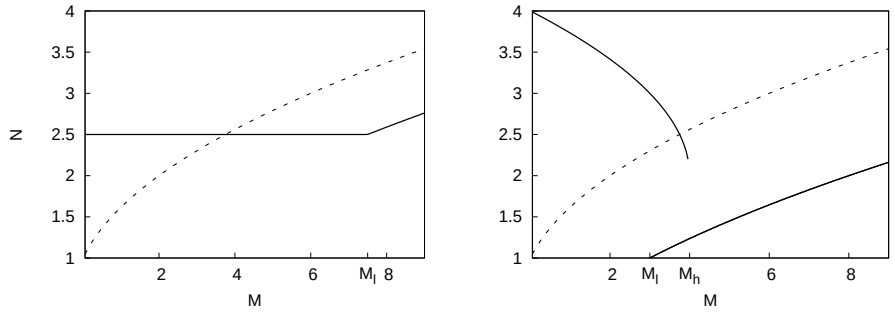
Figure 3.2 shows the equilibrium population size, allele frequency and average fitness for haploids and diploids. Haploid equilibria are similar to the weak selection case discussed in the previous section: the adaptive and maladaptive states are separated by a critical rate M_l, just as in Haldane's mainland-island model, see figures 3.2a and 3.2b. The key difference is demographic: the average fitness becomes negative for migration rates above rate M^{**} and the population becomes a sink (Pulliam 1988, 2000), see Fig. 3.2c,

$$M^{**} = \frac{1}{4}(2+s)^2. \tag{3.10}$$

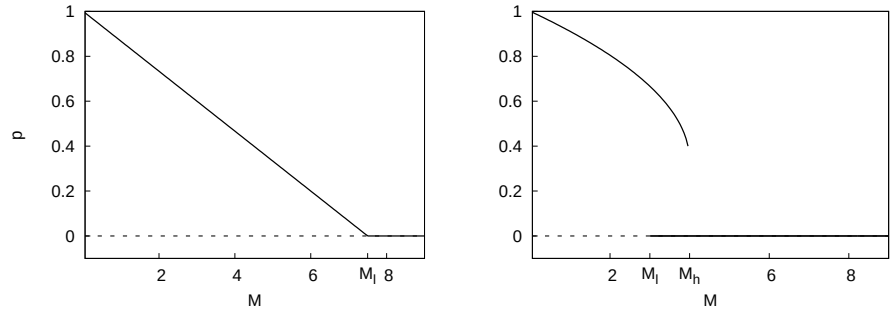
If the population is polymorphic, it can still become maladapted for migration rates above M^*, given by Eq. 3.9

The diploid equilibria, however, are markedly different. An upper threshold migration rate, M_h, appears in addition to Haldane's critical rate M_l. The latter arises due to the balance between migration and selection that is maintained at equilibrium; similarly, the former is due to the balance between growth and selection (more precisely selective deaths i.e. the elimination of deleterious alleles) that must also be balanced at equilibrium. For an intermediate range of migration rates above rate M_l and below the upper critical rate M_h, the high and low states are simultaneously stable, see figures 3.2a and 3.2b. In this case, the final equilibrium depends on the initial state of the population. For high migration rates above M_h, only the monomorphic (maladapted) state is stable. Additionally, the average growth rate associated with the monomorphic state is negative and the population is a sink, similar to the haploid case above, see Fig. 3.2c.

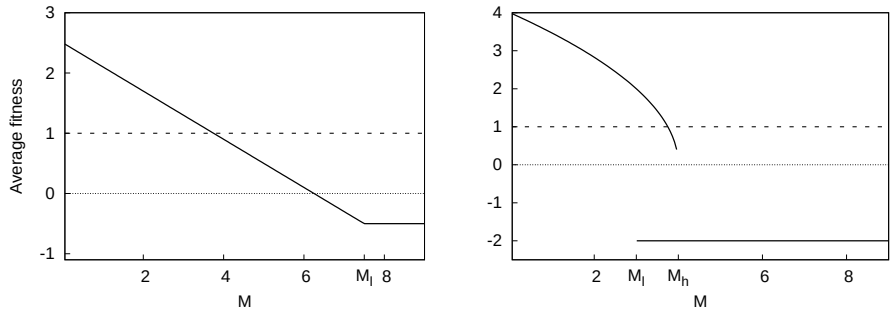
Diploid bistability arises because of the ploidy effect discussed earlier: a diploid locus carries a greater migration load than a haploid locus and therefore the population size reduces to the carrying capacity and the growth rate becomes negative. Once again, the population becomes a sink. Negative growth rates



(a) *Population size*



(b) *Allele frequency*



(c) *Average fitness*

Figure 3.2 *Single locus deterministic equilibria when selection dominates growth, one locus: equilibrium (a) population size, (b) allele frequency and average fitness for haploids (left) and diploids (right) with* $s = 3$ *and* $\hat{p} = 0$*. The stable polymorphic (upper curves) and monomorphic equilibria (lower curves) are shown. The critical migration rates are:* $M_l = 7.5$ *for haploids,* $M_l = 3$*,* $M_h = 4$ *for diploids. For* $M < M^* = 3.75$ *adaptation occurs, and for* $M > M^{**} = 6.25$ *haploid fitness becomes negative. Neutral equilibria are also shown (dashed lines).*

and population sizes at or below the intrinsic carrying capacity seem to be requisites for bistability. In contrast to the weak selection case, the high steady state remains stable for migration rates higher than the critical rate M_l; the local allele persists under relatively larger migration loads than when growth is stronger than selection. However, this depends on the initial state of the population: if initially the population lies within the domain of attraction of the

polymorphic (monomorphic) state then the local population will be polymorphic (monomorphic) at equilibrium for migration rates $M < M_h$ ($M > M_l$) (Thompson & Stewart 1986, Strogatz 2004). A linear stability analysis of both states is presented in App. A. The bistability is characterised in greater detail in the next section, where we also determine the critical rate M_h.

Irrespective of ploidy and migration, the population size is at least as large as the intrinsic carrying capacity, even when the island turns into a demographic sink. Therefore, once again, maladaptation does not incur a demographic penalty. Once the local allele is lost, migration sustains population numbers. Note that the population is actually a pseudo sink (Watkinson & Sutherland 1995), since without migration it would maximise fitness and attain large population numbers, as per Fisher's Fundamental Theorem of Natural Selection (Fisher 1958).

We have considered the effect of relatively strong and weak growth when gene flow is totally maladaptive. We have found that source-sink structure emerges and, for diploids, bistability for an intermediate range of rates, when growth is weak relative to selection. Next we take a closer look at the diploid bistability, before moving on to consider polymorphic gene flow.

Bistability and critical migration rates

Bistability occurs when the average fitness falls below zero. We define the 'bistable range' as the range of rates bounded by critical rates M_l and M_h. Bistability implies that historical events have a bearing on the population's fate; in addition to the relative strengths of migration, selection and growth, the population's initial state prior to receiving immigrants also matters.

Rate M_l separates weak/strong migration (relative to selection) when growth is strong and is determined by the balance between selection and gene flow. The upper critical rate M_h emerges when growth is weak relative to selection. An analytical expression for M_h is found using quartics $F_x = Q_x(x, \hat{p}) + x$, where $Q_x(x, \hat{p})$ is given by Eq. 3.5. When $M = M_h$, curves $y = F_N(N)$ and $y = F_p(p)$ are tangential to lines $y = N$ and $y = p$ respectively. This gives two conditions that rate M_h must satisfy, namely $F_x(x) = x$ and $F'_x(x) = 1$, (see App. A for details). The diploid critical rates are,

$$M_l = s, \tag{3.11a}$$

$$M_h = \frac{1}{4}\bar{r}^2_{max}. \tag{3.11b}$$

Both rates depend on selection. The stronger selection is, the greater the 'width' of the bistable range, which we define to be $\triangle M = M_h - M_l$. For $s \leq 1$, the upper critical rate falls below M_l and bistability disappears, as shown in Fig. 3.2.

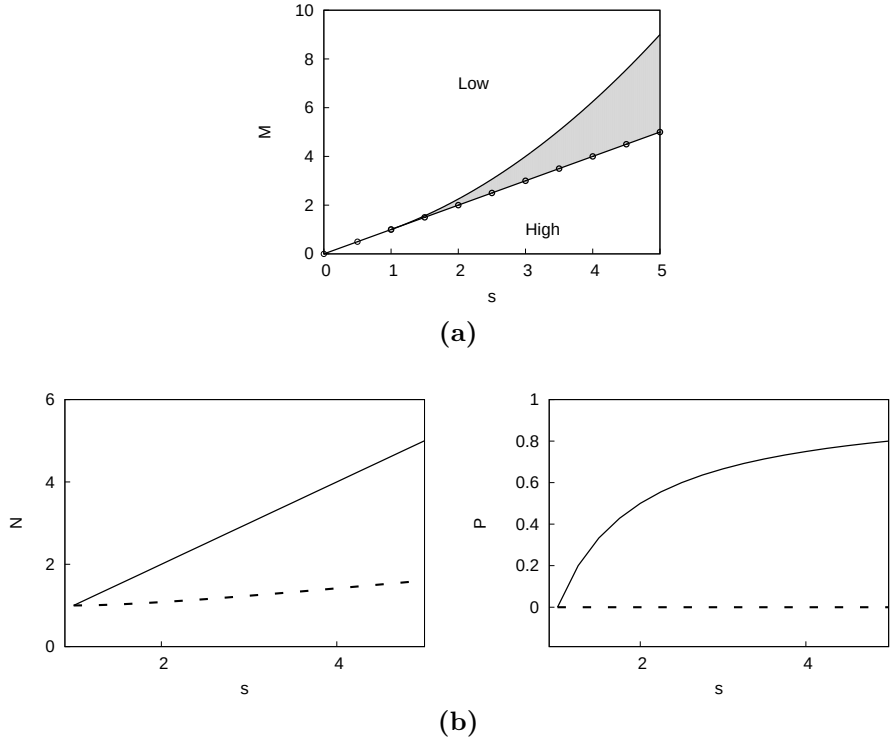


Figure 3.3 *Single locus critical migration rates and corresponding equilibria: (a) M_h (solid line) and M_l (with circles) for varying s. Above M_h the low state is stable, below M_l the low state is stable. Bistability occurs for points (s, M) that lie within the shaded region between the curves. For $s < 1$ diploid bistability disappears. (b) equilibrium population size (left) and allele frequency (right) for critical migration rates M_l (solid line) and M_h (dashed line).*

The bifurcation curve in Fig. 3.3a shows the changing bistable width (shaded region) for increasing selection. In Fig. 3.3b the equilibrium population size and allele frequency for critical rates M_l and M_h are shown. At rate M_l, the population size is $N = \overline{r}_{max}$ and the allele frequency is $p = 1 - 1/\overline{r}_{max}$. The population size reduces to $N = \overline{r}_{max}/2$ at migration rate M_h, with $p = 0$. The high state population size for M just below M_h is the largest size attainable by the local population whilst the favoured allele persists. This is local adaptation under the largest migration load experienced by the population for a given selection strength s. Furthermore, for $s \geq 2$, the island population is biased towards the local allele for migration rates $M \leq M_h$, when the polymorphic allele frequency satisfies $p_h > \frac{1}{2}$.

The importance of historical events

Characteristic of nonlinear dynamical systems with multiple stable states, there is hysteresis: the population 'jumps' between the adaptive and maladaptive states at the critical migration rates (Thompson & Stewart 1986, Strogatz 2004).

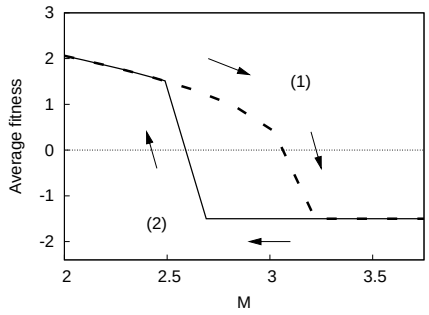


Figure 3.4 *Hysteresis in the average diploid fitness $\overline{r}$ for varying migration with $s = 2.5$, $\hat{p} = 0$. The population 'jumps' between states at critical rates M_l and M_h. The arrows indicate two different paths: (1) the island is initially polymorphic and migration is increased from a low to a high rate and (2) the island is initially monomorphic and migration is reduced from a high rate.*

In Fig. 3.4 the average diploid fitness is shown. Suppose the population is initially polymorphic and migration is low, $M < M_l$ (path 1, top arrow in figure). Migration begins to increase. The population remains polymorphic until rate M_h, when the migration load becomes too large, the local the allele is suddenly lost and the island becomes monomorphic. Now, suppose the migration rate is reduced (path 2, bottom arrow in figure). The population remains monomorphic until rate M_l, when it suddenly becomes polymorphic. If the local allele is lost from the mainland, $\hat{p} = 0$, hysteresis makes little biological sense: once the allele is lost from the island it cannot spontaneously reappear (without mutation). However, as we show in the next section, if the local allele is rare on the mainland, $\hat{p} \sim 0$, the island can become polymorphic once again at M_l and hysteresis is biologically interesting.

We have investigated the effect of total maladaptation on the island population and found that local adaptation depends not only on the relative strengths of selection, migration and growth, but on ploidy also. For migration rates that lie within the bistable range, local adaptation may also depend on the state of the population prior to the onset of gene flow. Next we examine equilibria for polymorphic gene flow, in particular, for when the local allele is rare, $p \sim 0$.

3.3.5 Polymorphic gene flow

Suppose the migrant gene pool is polymorphic. Exact steady state solutions can be obtained from the quartics in Eq. 3.5 but the expressions are unwieldy. Instead we determine equilibria by numerically solving the dynamics using MATHEMATICA (Wolfram 1999). The implicit steady state solutions are of course

$$N = \frac{\overline{r} + \sqrt{\overline{r}^2 + 4M}}{2}, \tag{3.12a}$$

$$p = \frac{(Ns - M) \pm \sqrt{(Ns - M)^2 + 4NsM\hat{p}}}{2Ns}. \tag{3.12b}$$

Figure 3.5 shows the equilibrium diploid population size and allele frequency for different degrees of migrant polymorphism. When the local allele is rare on the mainland, $\hat{p} \sim 0$, for example $\hat{p} = 0.002$, the situation is similar to the case of totally maladaptive gene flow and two stable equilibria emerge, see figures 3.5a and 3.5b.

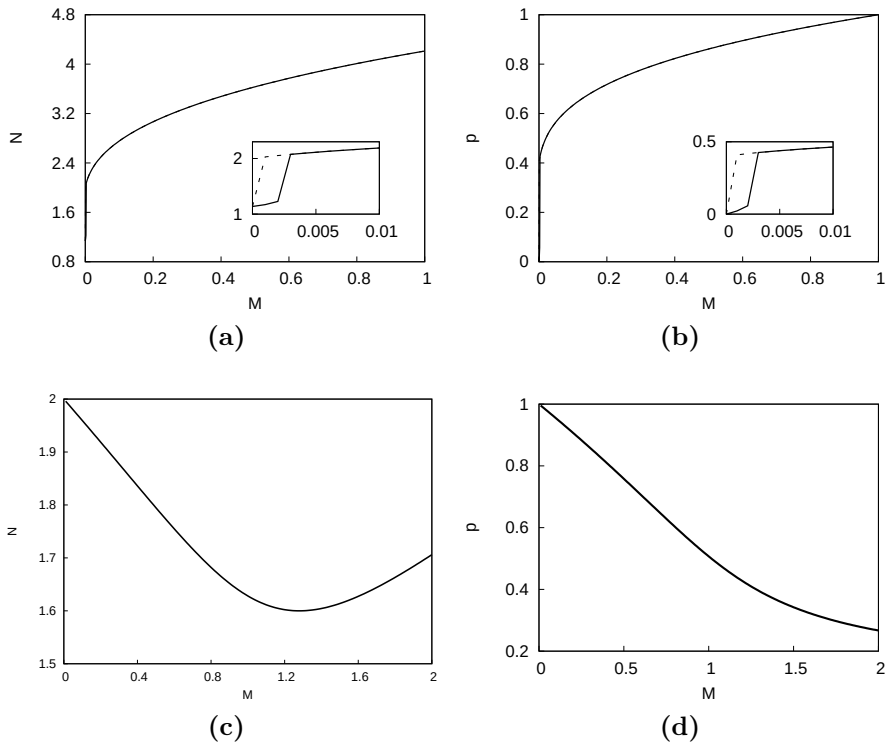


Figure 3.5 *Single locus equilibria for polymorphic gene flow: diploid equilibrium population size (left) and allele frequency (right) are shown. Top row, $\hat{p}$ is varied with $s = 2.5$ and $M = 3$. Bottom row, migration rate M is varied with $s = 1$ and $\hat{p} = 0.1$. When the local allele is rare, $\hat{p} \sim 0$, bistable equilibria emerge (top row, inset).*

As $\hat{p}$ increases, bistability disappears and the population size and allele frequency smoothly decrease with increasing migration. Once migration is high enough to compensate for maladaptation, the population size begins to increase, see figures 3.5c and 3.5d. Notice that the population turnaround occurs for $p > \hat{p}$, before the island is completely swamped by migrant genes. The greater the migrant polymorphism, or the more common the local allele is on the island, the smaller the migration load since the smaller the initial heterogeneity between mainland and island. As would be expected under less maladaptive gene flow, the population size is larger and the allele frequency is higher.

We have come to the end of our examination of the single-locus model. We have found that when the local allele is rare or lost on the mainland, local adaptation requires either weak selection and migration relative to strong growth, or for weak growth, strong selection and low to intermediate migration. Furthermore, the feedback between demography, adaptation and migration can turn the mainland-island system into a source-sink system when selection is stronger than growth and introduces an upper critical threshold above which adaptation is not possible.

We now turn to polygenic evolution and first investigate the multilocus model, before moving on to the quantitative trait model.

3.4 Multilocus case

In this section we investigate the consequences for local adaptation when multilocus evolution is coupled to population dynamics. We first determine the average fitness, which is a simple extension of the single locus average fitness. We then examine two distinct cases: when loci are symmetric (they have the same effect on trait), and when they are asymmetric. The former is known as the 'symmetric model' (Barton 1992). It is a simple extension of the single locus case and can be solved analytically. The asymmetric model, on the other hand, quickly becomes difficult to analyse exactly for more than a few loci. We will therefore treat the asymmetric case as a perturbation from the corresponding symmetric case. We again consider total maladaptive gene flow and polymorphic gene flow. We will find that though the multilocus results are qualitatively similar to the single locus case, important differences arise. Furthermore, asymmetries across loci can dramatically disturb bistability at equilibrium.

3.4.1 Trait average and variance

As outlined in §3.2.1, we assume n loci contribute to the trait and that loci evolve independently according to the dynamical equation Eq. 3.1b. Suppose the small contribution of locus j to the trait is $\frac{1}{2}\alpha_j$ ($-\frac{1}{2}\alpha_j$) for each gene that expresses the favoured (deleterious) allele. The sum of contributions over loci equals the final value of the trait. For n diploid ($\eta = 1$) or haploid ($\eta = \frac{1}{2}$) loci, we define the trait, z', to be

$$z' = \begin{cases} \sum_{j=1}^{n} \alpha_j (I_{j1} + I_{j2} - 1) & \eta = 1, \\ \sum_{j=1}^{n} \alpha_j (I_{j1} - \frac{1}{2}) & \eta = \frac{1}{2}, \end{cases} \tag{3.13}$$

where indicator variable I_{jk} is '1' if the k^{th} copy of the gene at the j^{th} locus expresses the favoured allele and '0' otherwise. The trait mean z and variance v_z are

$$z = \eta \sum_j \alpha_j (2p_j - 1), \tag{3.14a}$$

$$v_z = 2\eta \sum_j \alpha_j^2 p_j (1 - p_j). \tag{3.14b}$$

To determine the average fitness $\overline{r}$, we follow the procedure in §3.3 (but using scaled variables). In the equivalent discrete population, we take the fitness to be an exponential function since selection is directional; therefore the Wrightian fitness is $W(z') = e^{\beta z'}$, where β is the selection gradient (the slope of the graph of fitness against trait; Barton et al. 2007, Ch. 15). The average Wrightian fitness is then approximately

$$\begin{aligned} \overline{W} &= \overline{e^{\beta z'}} \\ &\approx 1 + \beta z, \end{aligned} \tag{3.15}$$

where z is the trait mean. Since $\ln(\overline{e^{\beta f(x)}}) \approx \overline{\beta f(x)}$ when β is small, the average fitness is

$$\overline{r} = 1 + \beta z. \tag{3.16}$$

The selection pressure on locus j is then $\beta\alpha_j$. Each locus therefore evolves

according to

$$\frac{dp_j}{dt} = p_j(1-p_j)\beta\alpha_j + \frac{M}{N}(\hat{p}_j - p_j), \tag{3.17}$$

which is just Eq. 3.1b with Eq. 3.16 and Eq. 3.14a substituted in. The equilibrium behaviour of the population is described by the simultaneous solution of $n+1$ dynamical equations. In general, exact analytical solutions are not possible. Note that the specific form of the trait does not matter. To determine the selection pressure on each locus, we need only specify how the average fitness varies with average trait, i.e. the selective gradient given by $\partial_z \overline{r} = \beta$, and the contribution of the j^{th} locus to the trait, α_j. However, for the sake of clarity we have specified a particular form for the trait.

However, if loci are identical, exact analysis is possible in the limit of total maladaptive gene flow, $\hat{p}_j = 0$. Under symmetry, the number of degrees of freedom reduce to just two, the population size and the symmetric allele frequency. As we will show in the next section, the qualitative results are therefore similar to the one locus case. However, important differences arise due to the coupling between selection and growth through the effective selection coefficient.

3.4.2 Symmetric loci

The allelic effect, α, and migrant frequency, $\hat{p}$, are identical across symmetric loci, and each locus has the same equilibrium allele frequency (by definition). The multilocus model can therefore be reduced to an equivalent single locus model. Setting $\alpha_j = 1$ and $s = \beta$ in Eq. 3.17, the symmetric multilocus model has the same dynamics as the single locus case in Eq. 3.4, except that the average fitness in Eq. 3.4a now depends on n loci,

$$\overline{r} = 1 + ns\eta(2p-1). \tag{3.18}$$

The effect of n loci is to increase the sensitivity of growth to adaptation by 'inflating' the effective selection coefficient, which is now $ns\eta$. In fact, the single locus model is a special case of the symmetric model. The selection strength per locus is still s.

We analyse the symmetric case in much the same way as the single locus model; the hard work has already been done, and the single-locus results are easily extended to account for n loci. We again consider and contrast the equilibrium behaviour of the island in the case of totally maladaptive gene flow ($\hat{p} = 0$) and polymorphic gene flow ($\hat{p} > 0$).

Totally maladaptive gene flow revisited

What effect do symmetric loci have on local adaptation and demography when the local allele is lost on the mainland? The equilibrium states of N and p were given in §3.3.4. These are Eq. 3.7 and Eq. 3.6 respectively, but now with the average fitness as given by Eq. 3.18.

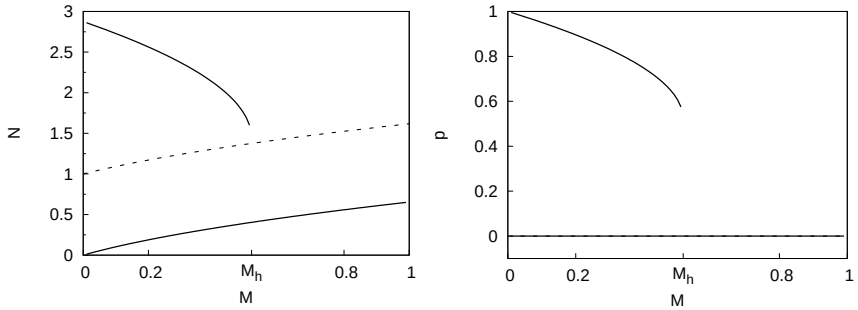


Figure 3.6 *Deterministic equilibria for 5 symmetric haploid loci: equilibrium population size (left) and allele frequency (right) with* $s = \frac{3}{4}$ *and* $\hat{p} = 0$*. The upper critical migration rate is* $M_h = \frac{529}{1024}$*. There is no biologically valid lower limit to the bistable range (i.e.* $M_l < 0$*). Neutral results are also shown (dashed lines).*

In Fig. 3.6 the equilibrium population size and symmetric allele frequency are shown for five haploid loci. Similar to the single-locus results (see Fig. 3.2), the island is either polymorphic with a positive average growth rate, or it is fixed for the deleterious allele and becomes a (pseudo) sink. The states are bistable for a range of migration rates. Above an upper critical rate M_h, the favoured allele is lost and the local population is maladapted.

Aside from these common features, the equilibria in Fig. 3.6 also highlight the differences between the single-locus and symmetric results. These differences arise solely because of the 'inflated' effective selection coefficient, $ns\eta$: growth becomes increasingly sensitive to adaptation with large loci numbers. It takes fewer deleterious alleles, hence a lower migration rate, for the average fitness to fall below zero and hence for the local population to become a sink.

Weak selection

In the single locus case, the polymorphic and monomorphic states are bistable for a diploid locus when the effective selection coefficient satisfies $s\eta > 1$. For diploids, there is no distinction between the effective selection coefficient and the selection coefficient per locus, so s has to be greater than 1. For symmetric loci,

the more general condition for bistability is $sn\eta > 1$. What this means is that the selection strength per locus, s, can be weaker than the intrinsic growth rate and bistability will still exist, provided locus number n satisfies $n > \lfloor \frac{1}{s\eta} \rfloor$, where $\lfloor x \rfloor$ is the floor of x. Since the effective strength of selection is spread across n loci, $s \approx \frac{1}{n}$, a relatively small selective advantage could have a large effect on demography if many loci experience the same selective advantage.

The importance of ploidy

In the single locus case, because a diploid locus experiences a greater migration load per locus than a haploid locus, bistability only exists for diploids. In the symmetric case, haploid loci can also be bistable, provided $sn\eta > 1$. It would seem the ploidy parameter has been rendered obsolete, since it features in the dynamics only as one part of the product $n\eta$. There is indeed an equivalence now between diploids and haploids: for a given selection strength per locus s, n diploid loci are indistinguishable from $2n$ haploid loci in our model. However, for an odd number of haploid loci, there is no diploid equivalent. Note that this equivalence is merely a mathematical observation and has no biological relevance since ploidy is generally preserved throughout an individual's life history (with the exception of meiosis).

Extreme population sizes

The equilibrium population size is at least as large as the carrying capacity size, irrespective of the parameters, in the single locus case. In the symmetric case however, since the effective selection coefficient can be large if n is large, even if selection is weak, the difference between the maximum and minimum average fitnesses increases, thus increasing the difference between the high and low state population sizes. As a result, population size in the maladapted state can fall far below the intrinsic carrying capacity, ($N < 1$). For this to happen, the inequality $s\eta(1-n)(s\eta+1) < 0$ must be satisfied. This is only possible if more than one locus contributes to trait, $n > 1$. If loci numbers are large, say $n > 20$, the high and low state population sizes become 'polarised', taking on extremely large and extremely small values respectively.

Critical migration rates

We can determine the critical migration rates M_l and M_h for n symmetric loci in the same we did for one locus, see Eq. 3.8 and §3.3.4. Rate M_l is found by rearranging $N_l = M_l/s$. To find M_h, we form quartics $Q_x(x, \hat{p})$ (where $x = N$ or p) similar to those in Eq. 3.5 (see §3.3.3) but using the symmetric dynamics. We then form the functions $F_x(x) = Q_x(x, \hat{p}) - x$ and use the fact that when $M = M_h$, the conditions $F_x(x) = 0$ and $F'_x(x) = 1$ are satisfied. The critical migration rates for the symmetric model are

$$M_l = s(\overline{r}_{min} + s), \tag{3.19a}$$

$$M_h = \frac{\overline{r}^2_{max}}{4(2n\eta - 1)}. \tag{3.19b}$$

where $\overline{r}_{min} = 1 - sn\eta$ and $\overline{r}_{max} = 1 + sn\eta$. We recover the one locus expressions with $n = 1$, cf. Eq. 3.11. Unlike in the single locus case, M_l can be negative if $s > 1/n\eta + 1$. That is, if n is large or s is sufficiently weak, the lower limit of the bistable range becomes unbiological and hence disappears; the high and low states are bistable for almost zero migration. What this means is that if the population lies initially within the domain of the low state, i.e. it is initially maladapted, it will have little chance of adapting, even under weak migration. The critical migration rates are shown in figures 3.7a and 3.7b for two and four loci respectively.

The bistable width increases for large locus numbers and/or strong selection. If the population is initially within the domain of polymorphic state, it can withstand maladaptive gene flow for high migration rates when selection is relatively weak, due to the cumulative effect of selection on many loci.

The population sizes and allele frequencies at the critical migration rates are shown in figures 3.7c and 3.7d respectively. At M_l, the population size is 'inflated', many times the carrying capacity, whereas at M_h it remains close to zero for large s, i.e. it effectively collapses and is close to extinction. This 'collapse' is not seen in the single locus case, where the population size always remains at least as large as the intrinsic carrying capacity, i.e. $N \geq 1$. These 'polarised' population sizes can be seen clearly in the figures. As the number of loci that contribute to trait increase, changes in the population size become more sensitive to changes in the genetic make-up of the population due to the effective selection coefficient, $ns\eta$. Therefore, even if migration is low, it becomes possible for migration load to be

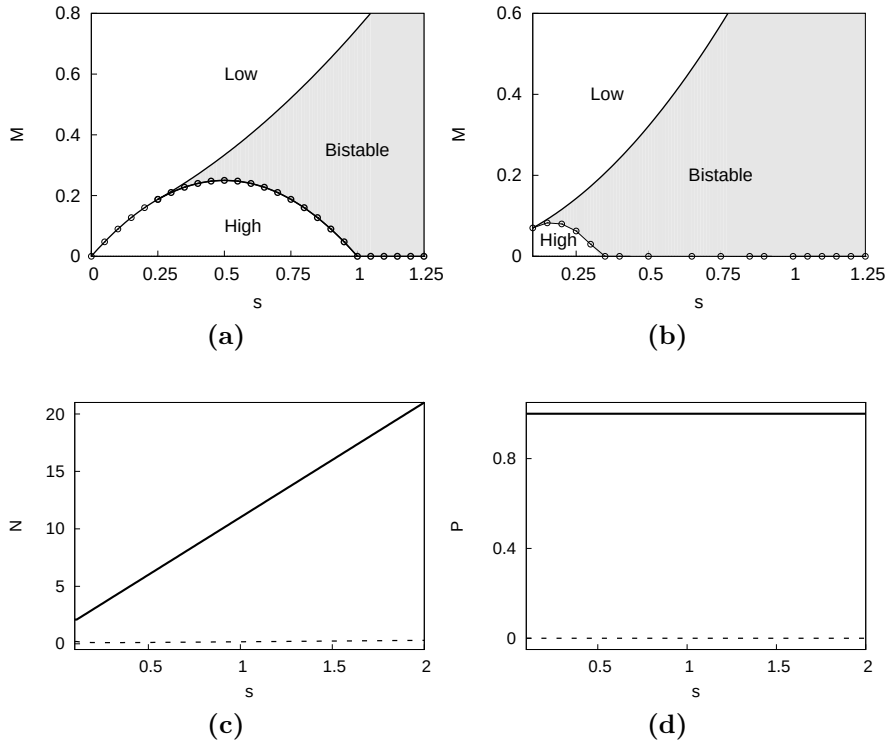


Figure 3.7 *Critical migration rates and equilibria for symmetric diploid loci: M_h and M_l (top row) and the critical population size and symmetric allele frequency (bottom row) for totally maladaptive gene flow. Top row, bifurcation curves showing M_l (lower curve) and M_h (upper curve) for (a) $n = 2$ and (b) $n = 4$. Where indicated, the high and low states are stable. Notice that the lower bound M_l almost disappears for $s < 1$ in (b). Bottom row, (c) the critical population size and (d) critical allele frequency at M_l (dashed curves) and M_h (solid curves) are shown for $n = 10$. The difference between the critical population sizes increases for increasing s, and the polymorphic allele frequency does not fall below $1/2$.*

high enough to cause the population size to collapse. The critical allele frequencies also appear 'polarised': the local allele is either lost or close to fixation. The larger the number of loci, the more polarised the critical states become. Again, local adaptation is possible under the greatest migration load for M just below rate M_h, that is, just before bistability collapses. Note that in contrast to the single locus case, in the high state, if n is large, the population is always well-adapted since the favoured allele is always the common allele, $p_h \gg 1/2$ (threshold M^* does not exist since rate M_l can 'disappear').

We have examined the symmetric model under totally maladaptive gene flow and found that the larger the number of loci the greater the difference between the high and low states. The penalty of maladaptation increases with increasing loci number, but local adaption is more robust against gene flow: however, this requires the local population to be in an adaptive state before

the onset of migration. The 'inflated' selection coefficient produces another important difference between the single locus case and symmetric case to do with polymorphic gene flow. We next explore the effect of symmetric loci on local adaptation when migrant gene flow is polymorphic.

Polymorphic gene flow revisited

In the single locus case, polymorphic gene flow destroys bistability for any amount of migrant polymorphism except when the local allele is rare. The analytical results are the same as the single locus case in Eq. 3.12, but with the multilocus average fitness given in Eq. 3.18.

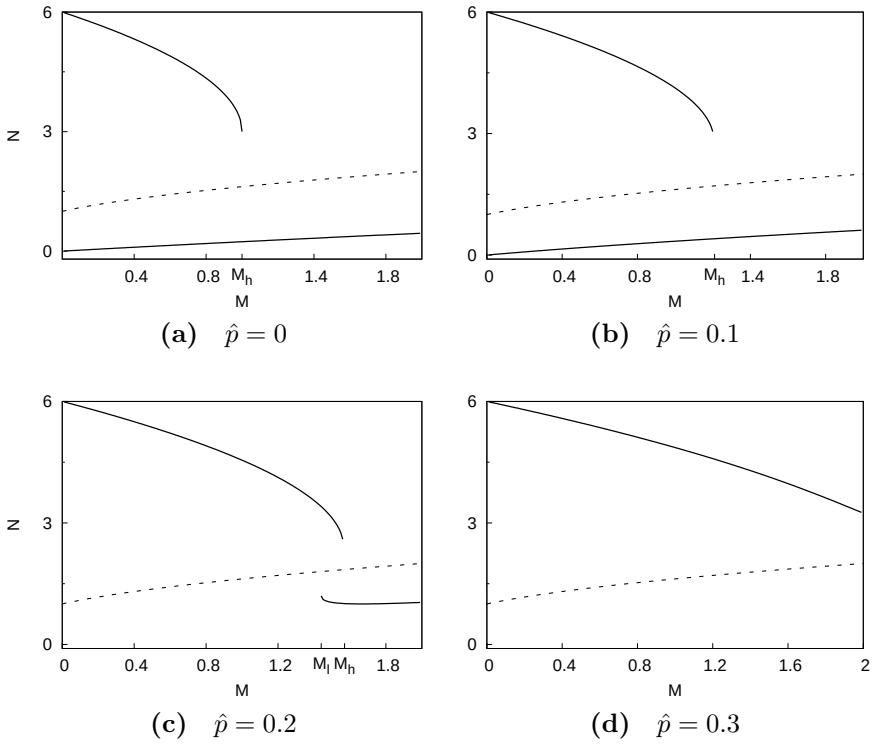


Figure 3.8 *Equilibrium population size for 5 symmetric, diploid loci for different migrant allele frequencies $\hat{p}$ with $s = 1$.*

Figure 3.8 shows the equilibria of different migrant frequencies $\hat{p}$ for five diploid loci. The steady states are bistable for polymorphic gene flow, in contrast to the single locus case, even when the local allele is not rare on the mainland, see Fig. 3.5. As gene flow becomes less maladaptive ($\hat{p}$ increases), the difference between the high and low state decreases, see figures 3.8b and 3.8c. Furthermore, the bistable width shrinks and both critical migration rates increase. This is because increasing migrant polymorphism decreases the migration load and

so the local population is less maladaptive. This is opposite to the effect of increasing loci number or selection strength. The local population can persist in the adaptive state for higher migration rates. However, $\hat{p}$ can still be high enough to destroy bistability for modest loci numbers and the population steadily becomes maladapted as migration increases, see Fig. 3.8d.

This concludes our discussion of the symmetric model. What happens when loci deviate from the simplifying assumptions of the symmetric case? In the next section we explore the how local adaptation is affected by asymmetric loci under joint demographic and adaptive evolution. In particular, we focus on when bistability exists between the high and low states.

3.4.3 Asymmetric loci

We have looked at the case where loci contribute identically to trait, $\alpha_j = \alpha$, and experience the same migrant frequency, $\hat{p}_j = \hat{p}$. Asymmetry across loci can arise in two ways. First, each locus can contribute a different amount to trait, that is, the allelic effect α_j is different for each locus. For example, for five loci the allelic effects may be $\alpha = \{0.01, 0.5, 0.2, 0.1, 0.3\}$. Different allelic effects result in some loci being more important to local adaptation than others with a smaller allelic effect on trait. Second, asymmetry may arise amongst loci if the migrant frequency at each locus is different, $\hat{p}_j$. Each asymmetric locus gives rise to a different steady state allele frequency (see Eq. 3.1b), in contrast to the symmetric loci case where loci have the same steady state frequency. Therefore, asymmetric equilibria consist of n equilibrium allele frequencies and the equilibrium population size. Exact analysis of asymmetric equilibria is intractable, since they depend on the simultaneous solution of $n + 1$ nonlinear, coupled equations. Numerical solutions, on the other hand, are easily obtainable using MATHEMATICA (Wolfram 1999).

Depending on the parameters, individual loci can have either one or two stable equilibrium allele frequencies, as in the single locus case. Each 'bistable locus' will have a different adaptive (high) and maladaptive (low) equilibrium allele frequency depending on the value of α_j and $\hat{p}_j$. Since loci are coupled to population size, for a given 'high' or 'low' population size the asymmetric equilibrium is stable only if *all* bistable loci are simultaneously at their respective high or low equilibrium allele frequency. Any mixture of high and low allele frequencies amongst bistable loci results in an unstable asymmetric equilibrium. Therefore, for asymmetric loci, as for symmetric loci, there can be at most two

stable equilibria, one adaptive and one maladaptive. In general, some loci will be fixed for the local allele, others will be lost and, depending on the initial state of the population, others will have either high or low levels of polymorphism.

A qualitative understanding of the asymmetric model is possible if we treat the asymmetric steady state equilibria as a perturbation from a symmetric equilibrium. We next examine the effect of small and large deviations in $\hat{p}$ and α on a symmetric equilibrium. In particular, we determine when these perturbations preserve bistability.

Perturbations from symmetry

Consider the n-locus symmetric equilibrium $\{N_0, p_0\}$ with symmetric migrant frequency $\hat{p}_0$ and symmetric additive effect α_0. Suppose that either $\hat{p}_j$ or α_j (but not both) is perturbed by an amount δ_j at locus j such that $\sum_{j=1}^{n} \delta_j = 0$ (all other parameters are kept constant). The latter constraint is always true since for an arbitrary pattern of deviations in either $\hat{p}$ or α, the symmetric migrant frequency can be defined as $\hat{p}_o = \frac{1}{n}\sum_j \hat{p}_j$, and similarly the symmetric average additive effect, $\alpha_o = \frac{1}{n}\sum_j \alpha_j$. If ϵ_j is the deviation of asymmetric allele frequency p'_j from the symmetric allele frequency p_0 at locus j, then $\sum_j \epsilon_j = 0$. For small ϵ_j, $\epsilon_j^2 \approx 0$, which means that the variance of deviations is approximately zero, $\sigma_\epsilon^2 = \overline{\epsilon^2} - (\overline{\epsilon})^2 \approx 0$. Taylor expanding p'_j about p_0 is

$$p'_j = p_0 + a_1(N')\epsilon_j + a_2(N')\epsilon_j^2 + a_3(N')\epsilon_j^3 + \ldots, \tag{3.20}$$

where coefficients $a_i(N')$ are functions of $\hat{p}_o$, α_0, η, M, s and the N' is the perturbed (asymmetric) equilibrium population size. But N' is dependent on the asymmetric average fitness. Substituting Eq. 3.20 into the average fitness Eq. 3.16 gives

$$\overline{r}' = \overline{r}_0 + \sum_{j=1}^{n} A_1\epsilon_j + A_2\epsilon_j^2 + A_3\epsilon_j^3 + \ldots \tag{3.21}$$

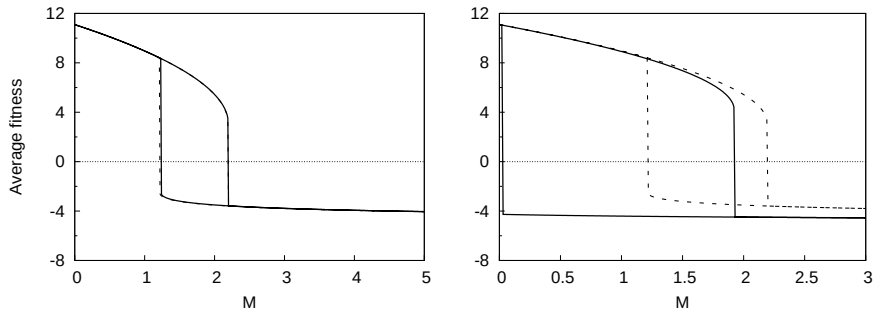
where $\overline{r}_0 = 1 + sn\eta(2p_0 - 1)$ is the symmetric average fitness and coefficients A_i depend on functions $a_i(N')$. Since the deviations ϵ_i sum to zero, the asymmetric average fitness only depends on second order (and higher) effects. If deviations are small, then $\overline{r}' \approx \overline{r}$. Therefore, the asymmetric equilibrium does not deviate significantly from the symmetric equilibrium when deviations ϵ_j are small enough that the variance of deviations σ_ϵ^2 is approximately zero. However, if $\sigma_\epsilon^2 > 0$, then second order (and higher) deviations become important.

In this case, it is difficult to predict what will happen, since the resultant asymmetric equilibrium depends crucially on the pattern of deviations across loci. However, only three outcomes are of interest: either the bistable width shrinks, expands or disappears altogether. Asymmetric equilibria cannot be bistable if the corresponding symmetric equilibrium is not bistable, since we have assumed that deviations sum to zero.

Below we provide numerical examples to illustrate the asymmetric equilibria that arise from small and large perturbations in the migrant frequency $\hat{p}$, followed by perturbations in the additive effects α.

Migrant frequency

Figure 3.9 shows the equilibrium average fitness for small and large perturbations in $\hat{p}$ ($\alpha = 1$ for all loci). What the expansion in Eq. 3.21 suggests, and what Fig. 3.9a clearly shows, is that symmetric equilibria are robust against small perturbations in $\hat{p}$, when $\sigma_\epsilon^2 \approx 0$.



(a) *small* $\sigma^2_{\epsilon_i}$ **(b)** *large* $\sigma^2_{\epsilon_i}$

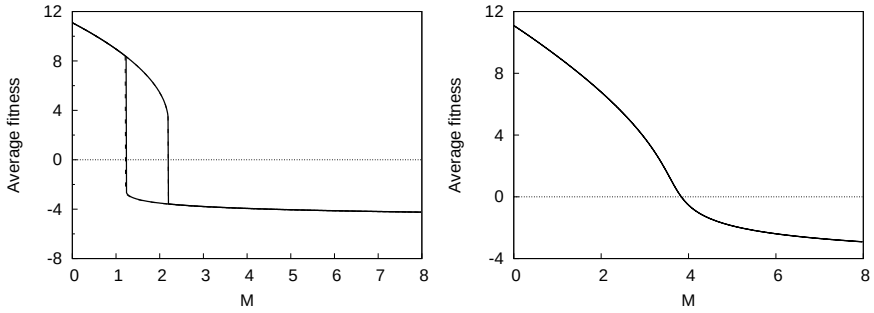
Figure 3.9 *Equilibrium average fitness* $\overline{r}'$ *for ten asymmetric diploid loci, with asymmetric migrant frequencies. (a) perturbations are small with variance* $\sigma^2_{\epsilon_i} = 0.006$. *The asymmetric average fitness (solid line) is barely distinguishable from the symmetric result (dashed line). (b) deviations are larger with a higher variance,* $\sigma^2_{\hat{p}} = 0.070$. *The symmetric parameters are* $\alpha_0 = 1.11$, $s = 0.800$, $\hat{p}_0 = 0.260$.

As shown in Fig. 3.9b, for relatively large deviations with higher variance, the asymmetric equilibrium differs significantly from the corresponding symmetric equilibrium. The bistable width increases because deviations in $\hat{p}$ are such that the overall migration load increases. Only a few large deviations are necessary to depress the average fitness of the population such that it is vulnerable to maladaptation at lower migration rates. Numerical investigations suggest that bistability is always preserved: if the symmetric equilibrium is bistable, then

the asymmetric equilibrium will also be bistable, albeit over a different range of migration rates.

Allelic effect

Deviations in α mean that selection varies across loci ($s = \beta\alpha$ and β is fixed). Again, if the variance of deviations is small, then the asymmetric equilibrium is barely distinguishable from the symmetric equilibrium, see Fig. 3.10a. Large deviations can the cause collapse of bistability, as shown in Fig. 3.10b. Notice that this increases the migration rate at which the island becomes a sink ($\overline{r}' < 0$), which suggests that large deviations in α can buffer the population against maladaptive gene flow.



(a) *small* σ_ϵ^2 **(b)** *large* σ_ϵ^2

Figure 3.10 *Equilibrium average fitness* $\overline{r}'$ *for ten, diploid loci with asymmetric additive effects* α_j*. (a) small deviations from the symmetric values, with variance* $\sigma_\epsilon^2 = 0.005$*. (b) relatively large deviations with a higher variance,* $\sigma_\epsilon^2 = 0.940$*. The symmetric average fitness (dashed line) has parameters* $s = 0.800$*,* $\alpha_0 = 1.11$*,* $\hat{p}_0 = 0.260$*.*

This can be understood by the fact that bistable loci behave in consort: either all occupy their respective adaptive states or all occupy their respective maladaptive states. If selection at one locus is large enough to destroy bistability at that locus, all other loci adjust to maintain stability. If the deviation at one locus is too large, stability cannot be maintained. Variations in $\hat{p}$ or α, by affecting bistability, can improve or diminish the possibility of local adaptation and persistence of the local allele, when compared to the corresponding symmetric case.

We have come to the end of our examination of the genetic model. We next investigate the joint evolution of population size and average trait under directional selection. We study this model to contrast the results with the genetic model, since both models experience the same dynamics and have the same trait.

Only, in the genetic model the allele frequencies are explicitly tracked, whereas in the trait-demographic model, only the trait mean matters. This leads to an entirely different picture of local adaptation, as we show next.

3.5 The quantitative trait model

Traits, such as body size or wing span, are the "foodstuff of evolution" (Lewontin 2004). In the explicit genetic model above, we modelled the selected trait as being expressed by either a single locus or multiple loci and followed the dynamics of individual allele frequencies together with the population size. However, the genetic basis of a trait, such as allelic effects or indeed the number of loci, are generally unknown. Understanding local adaptation in terms of the dynamics of the quantitative trait without recourse to the dynamics of the underlying genetics would therefore be ideal. This requires knowledge of the long-term dynamics of the trait, which in turn depends on making some assumption about the genetic architecture of the trait (Barton & Turelli 1989). The standard approach is to assume normality: the phenotypic distribution of many traits is approximately normal, or it can be made so in principle, by transforming to an appropriate scale (Lande 1976, 1979, Falconer & MacKay 1996, Wright 1968-78). Moreover, the transformation can be such that the trait variance is independent of the mean (Lande 1976). This observation together with the fact that little has been verified about what causes the underlying genetic variance (Barton & Turelli 1989), has led to the prevalent practice in evolutionary ecology of treating the additive genetic variance as approximately constant.

Normality should hold under Fisher's gene action model (Fisher 1918), in which it is assumed that many unlinked loci of small effect contribute to the trait (though a few loci can produce approximate phenotypic normality (Thoday & Thompson 1976). The Central Limit Theorem is invoked to justify normality when many loci segregate independently for the trait, though this is questionable in some cases (Turelli & Barton 1990). If the relation between genes and phenotype is assumed to be continuous and additive effects are small ($\beta\alpha$ is small), a Taylor expansion shows that this leads to normality (Barton & Turelli 1989). Therefore, we still make some assumptions about the underlying genotypic-phenotypic relationship. However, the key idea is that we assume that we cannot know the details of the underlying genetic dynamics of the trait, and/or that it is analytically intractable to solve for the dynamics of all the many loci and their nonrandom associations.

In the quantitative trait model, we consider the same mainland-island model as outlined in §3.2, but now under the joint dynamics of population size and the average trait, z, instead of the individual allele frequencies. We will show that under the constant variance assumption, the trait-demographic dynamics lead to entirely different results to the genetic model. This highlights the sensitivity of theoretical studies to the assumptions they rest upon.

3.5.1 Trait-demographic dynamics

We will now discuss the average trait dynamics. The dynamical equations for population size and the average trait are

$$\dot{N} = N(\overline{r}(z) - N) + M, \tag{3.22a}$$

$$\dot{z} = v_z s + \frac{M}{N}(\hat{z} - z). \tag{3.22b}$$

where $\hat{z}$ is the migrant trait mean. Equation 3.22a is the same equation for population size that appears in the genetic model (see Eq. 3.1a, Eq. 3.4a), save that $\overline{r}$ does not depend on the underlying frequencies, $\{p_j\}$, but instead on the value of z. Equation 3.22b can be understood as follows. The first term on the right-hand side is due to selection and is the selection pressure multiplied by the the trait variance (Bulmer 1980, Barton et al. 2007). Without any variation in the trait, selection would have no effect on the trait. We assume directional selection as before, so that the Wrightian fitness is an exponential function of the trait average, as discussed in 3.4.1. Therefore, the average fitness is $\overline{r} = 1 + sz$, and so the selection pressure is $\partial_z \overline{r} = s$. Although we have been able to derive $\overline{r}$ using the genetic dynamics, it is a valid fitness function for a trait under directional selection (Barton et al. 2007). The second term is due to gene flow and is therefore similar to the gene flow term in the allele frequency equation (see Eq. 3.1b, Eq. 3.17).

The dynamics do not form a closed set of equations: the rate of change of z depends on v_z, the dynamics of which depend on higher moments. To terminate this dependence and obtain a closed set of equations we assume the variance is constant (but see Turelli & Barton 1990). This is a key difference between the genetic model and the quantitative trait model. In the genetic model, the genetic variance, proportional to $p_j(1-p_j)$, evolves as p_j evolves. However, in the quantitative trait model we need make an approximation about v_z because we do not know what dynamics for v_z are likely to be, and moreover, we need to form a closed set of equations. By assuming v_z is constant, it can be eliminated from

the dynamics by rescaling selection and the trait mean: $Z = \frac{z}{\sqrt{v_z}}$ and $\hat{Z} = \frac{\hat{z}}{\sqrt{v_z}}$ and $S = s\sqrt{v_z}$.

3.5.2 Phenotypic equilibria

The dynamics admit two equilibria,

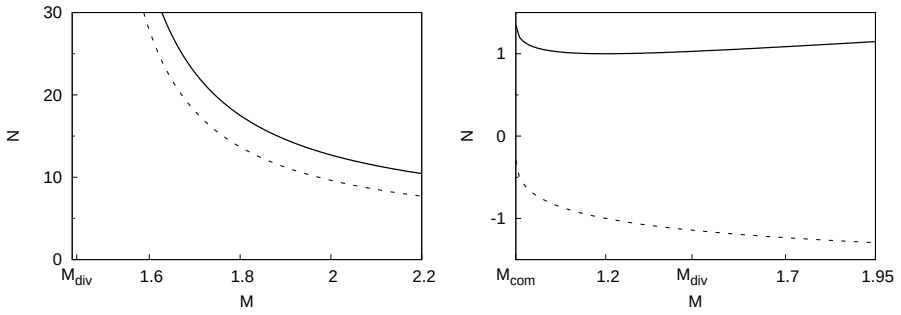
$$N_\pm = \frac{M\hat{r} \pm M\sqrt{(\hat{r})^2 + 4(M - S^2)}}{2(M - S^2)}, \tag{3.23a}$$

$$\begin{aligned} Z_\pm &= \hat{Z} + \frac{N_\pm S}{M} \\ &= \frac{S + \hat{Z}(2M - S^2) \pm S\sqrt{(\hat{r})^2 + 4(M - S^2)}}{2(M - S^2)}, \end{aligned} \tag{3.23b}$$

where here $\hat{r} = 1 + s\hat{z}$ is the average fitness of the migrants and is invariant under the variable transformation. How to interpret when local adaptation has occurred? Again, we use the genetic model to aid us. Suppose n loci contribute to the trait. Using the multilocus expression for the trait mean, Eq. 3.14a, we see that if Z is positive, then the allele frequency satisfies $p > 1/2$ at all loci (assuming all loci are identical); though the favoured allele is the more frequent allele on the mainland, gene flow is still imposes a large migrational load on the population. Conversely, if Z is negative, $p < 1/2$ at all loci, and therefore migrants are maladapted. Therefore, a positive average trait means adaptation has occurred.

There two types of equilibria, $\{N_+, Z_+\}$ and $\{N_-, Z_-\}$. At a critical migration rate $M_{\text{div}} = S^2$ all equilibria are indeterminate. Both equilibria become complex below a critical rate $M_{\text{com}} = S^2 - \frac{1}{4}(\hat{r})^2$. Linear stability analysis shows that state $\{N_+, Z_+\}$ is always stable when migration rate M is greater than M_{div}, otherwise it is a saddle node (Thompson & Stewart 1986). State $\{N_-, Z_-\}$ is always unstable.

The equilibrium population size and trait average of state $\{N_+, Z_+\}$ are shown in Fig. 3.11. If the favoured allele is the more frequent allele on the mainland, $\hat{Z} > 0$, the average growth rate never falls below zero, irrespective of migration and the migration load. At very small migration rates the population size diverges, with a correspondingly divergent trait average. As migration increases, the population size steadily reduces, in parallel with the diminishing average fitness, see Fig. 3.11a. This is a type of "migrational meltdown"; the population asymptotically approaches extinction with increasing migration. However, it is



(a) $\hat{Z} > 0$ **(b)** $\hat{Z} < 0$

Figure 3.11 *Deterministic trait-demographic equilibria. Equilibrium population size (solid lines) and average trait (dashed lines) of stable state* $\{N_+, Z_+\}$ *with* $S = 1.2$ *and (a)* $\hat{Z} = 2$ *and (b)* $\hat{Z} = -2$*. The critical migration rates are* $M_{com} < 0$ *for (a) and* $M_{div} = 1.44$ $M_{com} = 0.95$ *for (b).*

different from the original definition (Ronce & Kirkpatrick 2001), which critically depends on positive feedback between demography and evolution. Since the population size and average trait are unbounded at low migration rates, it is questionable whether this state has biological credibility.

For the more interesting case of maladapted migrants, $\hat{Z} < 0$, the island population is unable to adapt optimally and the average trait is always negative for nonzero migration. As migration increases, the island population decreases in size together with the average fitness. This is similar to the high state in the genetic model. This continues until the critical rate $M_l = S$ (Haldane 1931). At this rate, migration dominates and the population becomes a sink, similar to the low state in the genetic model. The population size begins to increase despite the falling average fitness due to the "rescue effect" of migration (Hanski 1999), see Fig. 3.11b. The critical migration rate M_l results from the balance between selection and migration and is the same critical migration rate found in in the single locus genetic model and soft selection models (see §2.2).

3.6 Summary and discussion

In this chapter we have investigated two deterministic models of local adaptation in a mainland-island metapopulation subject to (hard) directional selection, logistic density regulation and migration. In the first model, population size and n freely recombining loci evolve in consort. The second model is a trait-demographic model, in which genetic details are ignored and only the trait mean evolves alongside the population size.

The genetic model: we have solved the genetic model for a single locus and multiple symmetric loci. The basic interaction between selection and migration does not change with the inclusion of population dynamics: migration still 'swamps' the local population above a critical rate $M_l = s$ (for one diploid locus), in the 'low', maladaptive state. This is a classic result in population genetics Haldane (1931) (see Ch. 2). However, local adaptation is no longer guaranteed for migration rates below this threshold, and the threshold may disappear altogether. This is because the number of loci contributing to the trait matter: the strength of selection is spread across many loci which makes it easier for migration to overcome selection (Barton 1992).

Bistability: the inclusion of population dynamics means the constant population size varies in time, which extends the classical picture. The main results are: the appearance of a second critical rate, M_h, and the emergence of source-sink structure. Adaptation becomes possible for migration rates that are greater than M_l, that remain below M_h i.e. the population is able to withstand greater rates of maladaptive gene flow. In the multilocus case, it is possible for the critical rate M_l to disappear entirely, depriving the population of any chance of adapting in the face of migration, if it is not already well-adapted. If migrants are polymorphic, but the deleterious allele is the more frequent allele on the mainland, then the transition between the high and low states is smooth. In this case, neither critical rate is important. We have also derived critical rates for n symmetric loci. These rates are inversely proportional to the number of loci and proportional to selection, similar to the equivalent soft selection model (Barton 1992).

Once the average growth becomes zero or negative, the population becomes a (pseudo) sink and is dependent on migration to sustain its numbers ("rescue effect", Hanski 1999). Whether or not the local population achieves sink status depends on its state before the onset of migration: an initially small, poorly adapted population is more susceptible to sink status than if it were initially thriving and well-adapted. This could have implications for conservation: decreasing maladaptive gene flow may have little effect on a population due to its history.

Bistable range: we have shown that bistability can occur for a wide range of migration rates if selection is strong (for one locus) or alternatively selection is weak, in the multilocus case and the local allele is rare on the mainland.The transition between bistability and loss of the favoured allele becomes sharper as the number of loci increase. Furthermore, bistability is robust against small

perturbations from symmetry, but can be destroyed if the additive effects across loci are highly variable, for example polygenic traits with a few loci of large effect. It has been shown that the effect of loci can have different consequences on the conditions for genetic polymorphism (Kawecki 1995).

The scaled selection (per locus) has to be greater than the intrinsic growth rate, i.e. $s > 1$, for bistability to occur in the single locus case. The intrinsic growth rate is proportional to the strength of density dependence. In other words, density dependence must be weak for bistable equilibria to occur, as has been observed elsewhere (Tufto 2001). We expect that similar forms of density dependence, for example substituting logistic regulation with a hard population size cut-off, will produce qualitatively similar results. Since the intrinsic growth rate r_0 is scaled out, increasing the strength of density dependence does not change the results. However, the value of r_0 does affect the unscaled parameters: although the scaled values of selection may seem large, e.g. $s > 1$, the unscaled values, $s_u = sr_0$, will be small provided r_0 is small. We will discuss the biological plausibility of the parameter values in the next chapter and in the concluding chapter of the thesis, Ch. 6.

Bistability between the monomorphic and polymorphic states has been observed in similar single-locus studies (Holt & Gomulkiewicz 1997, Gomulkiewicz et al. 1999, Holt & Gomulkiewicz 2002). Our work is more general however: sink status is not imposed upon the local population and the local allele is not necessarily lost from the mainland. Furthermore, a detailed analysis of the critical rates M_l and M_h and the effect of multilocus evolution has not been carried out until now (see Barton (2001) for multilocus evolution in a spatial context).

In treating both haploid and diploid individuals in the same model, we have provided an explanation for why equilibrium bistability is absent from the haploid single locus model in Gomulkiewicz et al. (1999), which was not fully appreciated by the authors. Furthermore, the single locus haploid model can also be viewed as a diploid model of dominance. We appreciate, however, that in the genetic model it is more useful to view the ploidy parameter as controlling the sensitivity of population size to genetic evolution, since the distinction between haploids and diploids becomes unimportant in the multilocus case.

The phenotypic model: we have solved a simple quantitative trait model under the naive assumption of constant phenotypic variance. We have found that the basic condition for local adaptation under maladaptive gene flow holds: selection must exceed the rate of migration, as has been observed in the quantitative trait model studied by Tufto (2001). The model lacks the

more complex equilibrium structure of the genetic model because the dynamical equation for the average trait is linear as opposed to quadratic and therefore admits only one stable state. This explains why multiple stable equilibria have been observed in similar quantitative trait models with stabilising selection, which is quadratic in the trait mean (Kirkpatrick & Barton 1997, Barton 2001, Polechová et al. 2009, Tufto 2001, Holt et al. 2003). The markedly different outcomes of the phenotypic and genetic model could be explained as the difference between trait architecture: in the former, the trait consists of a large number of loci with small effects, but in the latter a (small) finite number of loci with large effects contribute to trait. Kawecki (1995) proposes a similar explanation for the apparent contradictory results regarding favourable conditions for polymorphism in fitness-sensitive models and single locus, multi-niche models.

Issues and future work: we have made the unrealistic assumption that associations between loci (linkage disequilibrium) are negligible (provided selection and migration are weak relative to recombination rates). However, migration introduces positive associations into the demic gene pool. If selection is strong, associations between loci will build up. Though nonrandom associations may be less likely in a growing population if growth is fast (Slatkin 1994). Since strong selection, migration and weak growth are likely to induce positive associations between loci, assuming linkage equilibrium is the 'worst case' scenario for adaptation under polygenic evolution. Little work has been done on linkage in genetic models with explicit population dynamics (see Barton 2001). This is an important issue in multilocus evolution and one that should be explored in future work, for example, by the study of a two-locus model in which linkage disequilibrium is allowed.

As both models have shown, the positive feedback between polygenic evolution and demography can lead to exceedingly small population sizes and low genetic variation. (We have equated population density with population number since we have assumed that the areal extent of the deme is constant. Though it has been argued that sink populations may contain more individuals than source populations despite lower population densities, if the areal extent of the sink is larger (Pulliam 1988).) Even if the population equilibrates to a large size, as it evolves it may evolve through a series of small sizes (see for example Gomulkiewicz & Holt 1995). Real populations may be vulnerable to extinction from, for example, (demographic) Allee effects, such as the inability to find a mate (Stephens et al. 1999, Keitt et al. 2000). Furthermore, the risk of inbreeding depression increases for small populations, which in turn could lead to elevated

extinction risk (Frankham & Ralls 1998, Saccheri et al. 1998, Nieminen et al. 2001, O'Grady et al. 2006). One way to incorporate extinction into the deterministic models would be to impose a low population size cut-off, below which extinction occurs. Alternatively, stochastic effects can lead to extinction (Lande et al. 2003). We explore this in the next chapter.

In summary, the deterministic models support conventional wisdom, whilst extending the classical picture of migration-selection balance and local adaptation: maladaptive gene flow can 'swamp' the demic population, but this is no longer a straight-forward case of the relative strengths of migration and selection. What these model have shown is that the relationship between the genetic and demographic dynamics of a population can be subtle, and ecological consequences can be significant when population growth is dependent on genetic parameters.

In the next chapter, we examine a stochastic analogue of the deterministic model.

Chapter 4

Stochastic mainland-island model

I now develop the central theme of the thesis and explore the stochastic effects of finite-sized populations, as described in §2.1.1. I do this by examining two continuous diffusion processes, one for each of the deterministic genetic and phenotypic models of the last chapter. I show that in the stochastic genetic model, when the local allele is rare and stochastic fluctuations are weak, the population can shift between two 'ecological zones': an adaptive zone and a maladaptive zone. The rate at which the populations escapes maladaptation is the rate of local adaptation; transition rates between these zones vary dramatically over a limited range of migration rates that critically depend on the number of loci. In contrast, the stochastic phenotypic model suggests that adaptation is not feasible under maladaptive gene flow, though population numbers may be high.

4.1 Previous studies

As discussed in §2.3.1, little work has been done on the effect of stochasticity on local adaptation in the species' range literature. Holt et al. (2005) studied quantitative trait evolution in a source-sink metapopulation subject to migration and stochastic events. The authors carried out individual-based simulations of a discrete, absolute sink population with nonoverlapping generations that received immigrants from a source population in mutation-selection balance. Individual loci were tracked in their simulations and free recombination was assumed. They found that under stabilising selection (quadratic dependence of average fitness on trait), maladaptive gene flow and density-dependent population regulation, a 'punctuated' pattern of niche evolution emerges in the sink: the population fluctuates about either a maladapted or well-adapted state during

long periods of stasis, between which the population rapidly transitions to the alternate state. The authors concluded that the addition of demographic and genetic stochasticity does not alter the basic, qualitative results drawn from corresponding deterministic models; demography constrains adaptation in sink environments and immigration can have positive and negative effects on niche evolution (Holt 2009).

Other studies of genetic and/or demographic drift are similarly simulation-based and have been discussed in Ch. 2, see §2.3.

These studies lack an accompanying theoretical treatment of the models they simulate. This is in part due to the difficulty in modelling stochastic eco-evolutionary dynamics. Moreover, species' range studies lack a general theory of local adaptation under stochastic evolution and growth when multiple loci contribute to trait. Here I develop such a theory by examining two distinct continuous stochastic models based on the deterministic models of the last chapter i.e. a multilocus genetic model and a quantitative trait-demographic model.

To develop a theoretical understanding of stochastic local adaptation, building on the deterministic mainland-island models of the previous chapter is a straightforward and obvious approach. The models I consider are the simplest that capture the essential deterministic and stochastic biological processes. They are also a natural continuation of earlier work; for example, the stochastic genetic model I develop extends Wright's island model (see §2.2.2), by explicitly including stochastic population dynamics. however, the models differ from earlier studies: I consider directional selection and multilocus evolution, as well as phenotypic evolution.

I employ the mathematical theory of diffusion processes and treat population size and allele frequencies as continuous Markov variables. As in the last chapter, I focus on the case where the local allele is rare on the mainland. I will show that in the limit of weak fluctuations, the population shows the same 'punctuated' pattern of adaptation and maladaptation as was observed by Holt et al. (2003) and that this is a feature of bimodality in the stationary distribution of states. The population is able to stochastically shift between two ecological 'zones', the species' niche and a maladaptive zone. My approach will be to treat the continuous models as "toy" models; the aim is to explore essential, qualitative features. Most of the chapter is dedicated to examining the genetic model, since it is the more complex model. I determine the parameters and special cases for which local adaptation is most likely, exploring the stationary and transient dynamics of the genetic model. There is no one result to single

out. Locus numbers and the relative strength of drift play a significant role in local adaptation. The stochastic phenotypic model behaves markedly different. However, the behaviour of the two models can be reconciled in the limit of large numbers. Finally, the continuous models can be viewed as diffusion approximations of a family of similar discrete models. I simulate a discrete model to test how well the genetic model approximates the discrete dynamics. As I show, the continuous theory is generally in good qualitative agreement with the discrete results.

The remainder of the chapter is structured as follows. In the first part of the chapter, Sec. 4.2, I outline the stochastic multilocus dynamics, before deriving the multilocus stationary distribution. I then examine the genetic model, first for one locus (§4.3) and then for multiple loci (§4.4). Concentrating on the rare allele case, I examine the probability of finding the population in an adaptive or maladaptive state and the stationary moments of the distribution for a single locus, analysing both the neutral and nonneutral cases, §4.3.2. Next, I define the 'bimodal range', which is analogous to the bistable width defined in the last chapter, and is the range of rates for which the stationary distribution has two ecological zones. I briefly consider polymorphic gene flow, before moving on to biologically relevant first-passage properties: the mean extinction time, fixation probability and the rate of transition between ecological zones, §4.3.4. To approximate the exact theoretical results, which in most cases can only be calculated numerically, I develop the local peak approximation scheme in §4.3.2. The multilocus analysis follows in a similar fashion, however I do not consider first-passage properties. In Sec. 4.4.4, I discuss the discrete model and simulation results and examine the validity of the continuous model as an approximation of the discrete model. Finally, in Sec. 4.5, I investigate the stochastic phenotypic model, before summarising the results of the chapter in the final discussion in §4.6.

4.2 The genetic model and dynamics

Consider the multilocus genetic mainland-island model from the previous chapter, except now the island population is finite-sized. All the assumptions and conditions of the deterministic model apply, except now the deme experiences genetic drift, which causes gene frequencies to fluctuate, and demographic stochasticity, which causes population sizes to fluctuate. We will refer to genetic drift and demographic stochasticity collectively as 'random drift' or just 'drift'.

By assuming the average generational changes are small, we can use a multivariate diffusion process to model the evolution of the population (Karlin & Taylor 1975*b*, Ewens 1979). Population size and allele frequency are now continuous Markov variables (their future states only depend on the current state of the population, Feller 1960). The stochastic dynamics can be written as stochastic differential equations (SDEs) of the Itô kind (Gardiner 1985, Karlin & Taylor 1975*b*).

Let $\mathbf{x}_u = \{N_u, \mathbf{p}\}$ (unscaled variables, see §3.2.1). Fluctuations only contribute to the conditional variance of generational change, $\mathrm{V}[\delta x_{i,u}|\mathbf{x}_u]$, and not to the average dynamics, where $\mathbf{x}_u$ is the current state of the population and $\delta x_{i,u}$ is the change over one generation in variable x_i (Karlin & Taylor 1975*b*). As was discussed in §2.2.2, the contribution from genetic drift is found from the dynamics of the Wright-Fisher model (Wright 1931, Crow & Kimura 1970). In the unscaled variables, this gives the conditional variance of generational change in the allele frequency of the j^{th} locus as $\mathrm{V}[\delta p_j|\mathbf{x}_u] = \frac{v_j}{2\eta N_u}$, where $v_j = \frac{p_j(1-p_j)}{2\eta}$ is the additive genetic variance at locus j and η is the ploidy parameter ($\eta = \frac{1}{2}$ for haploids, $\eta = 1$ for diploids). For population size, assuming the random birth and death of individuals follows a Poisson process (Murray 1989), then the conditional variance of generational change is $\mathrm{V}[\delta N_u|\mathbf{x}_u] = N_u$ (Lande et al. 2003, Murray 1989). Notice that migration does not contribute to fluctuations in N_u. This will be justified when we discuss the validity of treating the continuous model as the diffusion limit of discrete models (see §4.4.4). By once again rescaling the variables (see §3.3.2), a scaling parameter $\zeta = Kr_0$ arises due to the inclusion of the stochastic terms. The SDEs for the dynamics are

$$dN = \Big(\overline{r}N - N^2 + M\Big)dt + \sqrt{\frac{N}{\zeta}}dW_N, \tag{4.1a}$$

$$dp_j = \Big(v_j\frac{\partial\overline{r}}{\partial p_j} + \frac{M}{N}(\hat{p}_j - p_j)\Big)dt + \sqrt{\frac{v_j}{\zeta N}}dW_{p_j}. \tag{4.1b}$$

where $\overline{r} = 1 + \beta\sum_{j=1}^{n}\alpha_j(2p_j - 1)$. The differential dW_{x_i} is the increment of an independent Wiener process W_{x_i} with the properties $\langle W_{x_i}(t)\rangle = 0$, $\langle W_{x_i}(t)W_{x_j}(t_0)\rangle = (t - t_0)\delta_{ij}$ and $(dW_{x_i})^2 = dt$, where t_0 is some initial time (Karlin & Taylor 1975*b*). The terms multiplying dt represent the deterministic or average behaviour of the population; these are just the deterministic rate equations for N and p_j from the last Chapter (see Eq. 3.4a and Eq. 3.17). The terms multiplying dW_{x_i} represent the stochastic part of the dynamics. In Eq. 4.1a, this is just the square root of $\mathrm{V}[\delta N|\mathbf{x}]$, the variance of generational change in the

scaled population size N. Similarly, in Eq. 4.1b, the second term on the right is the variance of generational change of the allele frequency p_j in the scaled variables. Notice that the scaling parameter ζ only appears in the stochastic terms; in the deterministic dynamics the scaling parameter dropped out of the equations altogether (see §3.3.2), whereas it remains in the stochastic terms because of the square-root dependence of dW_{x_i} on dt. The scaling parameter 'tunes' the relative size of fluctuations. Therefore ζ takes on the role of the 'system size' or the 'large parameter' in the system (Van Kampen 2007). A large ζ means either the intrinsic carrying capacity K is large or the intrinsic growth rate is high (or both) and demographic stochasticity becomes negligible for very large populations. The diffusion process lies within the phase space domain defined by the intervals $N \in [0, \infty)$ and $p_j \in [0, 1]$. To define a unique diffusion process and a sensible stationary probability distribution, boundary conditions must be specified. This is discussed in the next section.

4.2.1 Stochastic equilibrium

The conditional, time-dependent probability distribution, $\Psi_n(\mathbf{x}, t|\mathbf{x}_0, t_0)$, is governed by a multivariate forward Fokker-Planck equation (FFPE) (Risken 1996). The one-dimensional FFPE was introduced in §2.2.2, when deriving the allele frequency distribution for Wright's island model. Using the shorthand $\Psi_n(\mathbf{x}, t) = \Psi_n(\mathbf{x}, t|\mathbf{x}_0, t_0)$, this is

$$\begin{aligned}\partial_t \Psi_n(\mathbf{x}, t) &= -\sum_i \partial x_i [A_i(\mathbf{x}) \Psi_n(\mathbf{x}, t)] + \frac{1}{2} \sum_{i,j} \partial x_i \partial x_j [B_{ij}(\mathbf{x}) \Psi_n(\mathbf{x}, t)] \\ &= -\nabla \cdot \mathbf{J}(\mathbf{x}, t), \end{aligned} \tag{4.2}$$

where $\mathbf{J}(\mathbf{x}, t)$ is the multivariate probability current, $\mathbf{A}$ is the deterministic field vector and $\mathbf{B}$ is the diffusion matrix (Gardiner 1985, Ch. 5). Vector $\mathbf{A}$ contains the directed 'forces' of growth and evolution and matrix $\mathbf{B}$ contains all the random contributions.

There is a correspondence between the SDEs and the FFPE that relates the coefficients of the 'dt' terms in Eq. 4.1 to the components A_i of $\mathbf{A}$, and similarly the coefficients of the 'dW' terms in Eq. 4.1 to the elements B_{ij} of $\mathbf{B}$ (Gardiner 1985, Ch. 4). It turns out that it is possible to write the deterministic dynamics in terms of a potential U, the components of $\mathbf{A}$ and $\mathbf{B}$ are

$$
\begin{aligned}
A_i &= N\partial_N U && \text{if } i = 1, \\
&= \frac{v_k}{N}\partial_{p_k} U && \text{if } i > 1,
\end{aligned}
\tag{4.3}
$$

and

$$
\begin{aligned}
B_{ij} &= \frac{N}{\zeta} && \text{if } i = j = 1\ , \\
&= \frac{v_k}{\zeta N} && \text{if } i = j > 1\ , \\
&= 0 && \text{otherwise.}
\end{aligned}
\tag{4.4}
$$

where the potential U is found to be,

$$
U = \overline{r}N - \frac{1}{2}N^2 + M\ln N + 2\eta M \sum_{j=1}^{n}\Big(\hat{p}_j \ln(p_j) + \hat{q}_j \ln(q_j)\Big) \tag{4.5}
$$

with $q_j = 1 - p_j$. The potential U will used to determine the stationary distribution and its definition will become clear below. In general, such a potential can only be defined under highly restrictive assumptions, such as assuming constant genotypic fitnesses.

The multilocus stationary distribution $\Psi_n(\mathbf{x})$ is found by setting Eq. 4.2 to zero and solving. The one-dimensional FFPE has been extensively studied and the methods to find its solution are numerous (Risken 1996). In higher dimensions however, finding the stationary distribution is often non-trivial. This is because it is possible to have circulating probability currents or limit cycles in higher dimensions. However, we can write down a solution to the homogeneous FFPE in terms of the potential U if we assume the probability current $\mathbf{J}(\infty)$ vanishes at equilibrium, provided the matrix B_{ij} has an inverse for all $\mathbf{x}$ (excluding the boundaries of the multidimensional diffusion interval). When this is the case, a set of potential conditions, $\{L_i\}$, related to the stationary distribution, can be defined (Gardiner 1985, pp. 146),

$$
\begin{aligned}
L_i(\mathbf{x}) &\equiv \sum_k (B_{ik})^{-1}\Big(2A_k - \sum_j \partial_{x_j} B_{kj}\Big) \\
&= 2\partial_{x_i} U - \frac{\partial_{x_i} B_{ii}}{B_{ii}} \\
&\equiv \partial_{x_i} \ln\Big(\Psi_n(\mathbf{x})\Big).
\end{aligned} \tag{4.6}
$$

From Eq. 4.6, we see that the potential condition L_i must be a gradient, which is true if $\partial_{x_j} L_i = \partial_{x_i} L_j$ (i.e. the curl of vector $\mathbf{L} = \{L_1, L_2, L_3, \ldots\}$ vanishes). We also see that the potential U is defined as the function that satisfies $\partial_{x_i} U = A_i/B_{ii}$ for all x_i. The stationary distribution then takes the form $\Psi_n(\mathbf{x}) = \exp \int^{\mathbf{x}} \mathbf{L}(\mathbf{x}') \cdot \mathrm{d}\mathbf{x}'$, similar to Wright's one-dimensional case discussed in §2.2.2. In terms of potential U given in Eq. 4.5, this gives

$$
\Psi_n = \frac{1}{Z_0}\frac{e^{2\zeta U}}{VN}, \tag{4.7a}
$$

$$
= \frac{1}{Z_0}\Big(N^{2M\zeta - 1} e^{-\zeta N^2}\Big)\Big(e^{2N\zeta \overline{r}}\Big)\Big(\prod_j^n p_j^{4M\zeta\eta\hat{p}_j - 1} q_j^{4M\zeta\eta\hat{q}_j - 1}\Big), \tag{4.7b}
$$

$$
= \frac{1}{Z_0} H(N) e^{2N\zeta\overline{r}} F_n(\mathbf{p}), \tag{4.7c}
$$

where $V = \prod_j v_j$ and $Z_0 = \int \Psi_n(\mathbf{x})\, d\mathbf{x}$ is the normalisation constant. The potential dynamics impose the condition that the probability current $\mathbf{J}(\infty)$ must be zero at stochastic equilibrium, which in turn implies the diffusion process obeys detailed balance (the property that transitions between any two states a and b satisfy

$$
\Psi_n(\mathbf{x}_a)\Gamma_{ab} = \Psi_n(\mathbf{x}_b)\Gamma_{ba}, \tag{4.8}
$$

where Γ_{ab} is the transition rate from state a to state b, Gardiner 1985). Since every solution of the FFPE converges to the same distribution in the long time limit (provided the diffusion matrix $\mathbf{B}$ is positive definite and does not contain any singularities, Risken 1996), Eq. 4.7 is the unique stationary solution, once the boundary conditions have been accounted for.

The stationary density can be separated into a product of three contributions: a 'demographic' term, $H(N)$, a 'genetic' term $F_n(\mathbf{p})$, and sandwiched between these two terms, a functional coupling between N and all p_j, $e^{2N\zeta\overline{r}}$. For constant population size $H(N) = 1$, $\Psi_n(\mathbf{x})$ reduces to Wright's multilocus allele frequency

distribution for the island model under fixed migration (gene flow) and directional selection (Wright 1931, cf. §2.4 in Ch. 2).

Provided migration is nonzero and $\hat{p} > 0$, the boundaries $N = 0$ and $p_j = 0, 1$ become reflecting so that if reached, the diffusing population is immediately 'reflected' to a state within the interior of the diffusion domain. The boundary at $N = \infty$ is never reached; it is a natural boundary (for boundary classifications see Karlin & Taylor (1975*b*), Gardiner (1985)). We denote the unnormalised stationary distribution as $\Phi_n = Z_0 \Psi_n$.

For $\Psi_n(\mathbf{x})$ to be a bona fide stationary probability density, the normalisation constant Z_0 must be finite and $\Psi_n(\mathbf{x}) \geq 0$ for all $\mathbf{x}$ within the phase space domain. This requires that the migration rate and migrant frequency $\hat{p}_j$ satisfy $M \geq \frac{1}{2}$ and $4M\eta\hat{p}_j \geq 1$, which prevents $\Psi_n(\mathbf{x})$ from diverging at the boundaries. These boundary constraints necessarily restrict the behaviour of the population. If we regard the continuous model as an approximation of a family of discrete models, we expect it to become an accurate approximation for discrete states away from the boundaries. This is because the diffusion description breaks down when the population consists of only a few individuals or a few copies of either allele (Ewens 1979). Nevertheless we will consider values of M and $\hat{p}_j$ for which $\Psi_n(\mathbf{x})$ diverges at the boundaries in order to assess how poorly the continuous theory approximates the behaviour of corresponding discrete models.

The multilocus stationary distribution has a rich variety of forms controlled by the relative strengths of selection, growth, migration and random drift. The potential U describes a multi-peaked, hyperdimensional surface that the population moves across as it evolves. Peaks in the 'potential landscape' coincide with peaks in the stationary probability distribution when fluctuations are small; peak maxima then correspond to stable deterministic states when fluctuations are removed. Potential 'valleys' connect neighbouring peaks. If fluctuations are large however, peaks in the potential U will not coincide with peaks in the stationary distribution; fluctuations move peak maxima in the stationary distribution away from the corresponding deterministic values. When demography and migration are ignored, this 'potential landscape' coincides with Wright's 'adaptive topography'. If the population size is large and kept constant, peaks Wright's 'adaptive topography' correspond to peaks in mean fitness. In this case, to reach a new adaptive peak the population would have to cross a valley of lower fitness before ascending the new peak (Wright 1931, 1932). In general however, mean fitness is not maximised when the population ascends a peak in the potential landscape (see e.g. Barton 1989).

With the stationary distribution in hand, before considering multilocus adaptation, we examine the single locus case. We first consider the stationary behaviour of the population under strong random drift, strong migration and strong selection, moving on to the main case of interest, local adaptation when the local allele is rare on the mainland.

4.3 Single locus

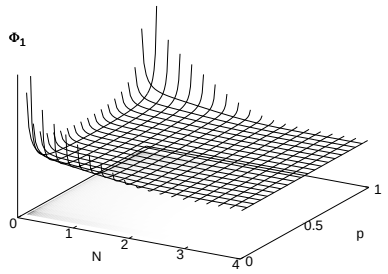
In this section we explore the effect of random drift on the population when fitness differences are caused by a single locus. Particularly, we wish to know the conditions under which the locally favoured allele can spread and the population can persist when the local allele is rare on the mainland, $\hat{p} \sim 0$, and when it is more common, $\hat{p} \gg 0$. We do this by finding the stationary properties of the population, by which we mean the average and variance of N and p, and the probability of finding the population near a particular state. We also look at the transient behaviour of the population, in particular the transition rates between alternate states. We will compare numerical calculations of exact results with theoretical approximations and simulation results from a discrete model. We consider a discrete model similar to the classic Wright-Fisher model but which includes fluctuating population size, the Modified Wright-Fisher Model (MWFM) (see Gomulkiewicz et al. (1999) for a similar model). The discrete model and simulation results are discussed in detail in §4.4.4.

First, we consider how the relative strengths of drift, migration and selection affect the stationary behaviour of the population. To do this we examine marginal stationary distributions of population size, $\phi_{N,1}$, and allele frequency, $\phi_{p,1}$, together with the bivariate single-locus stationary distribution, $\Psi_1(N, p)$.

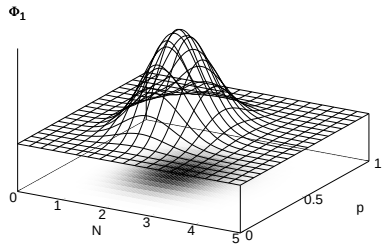
4.3.1 Random drift, migration and selection

We take the (scaled) selection coefficient to be $s = \beta$, setting $\alpha_j = 1$ in $\overline{r}$ (see §4.1b). The joint single-locus stationary distribution Ψ_1 has the same potential form as Eq. 4.7a, but without the product over n loci in F or the sum over loci in the mean growth rate $\overline{r}$.

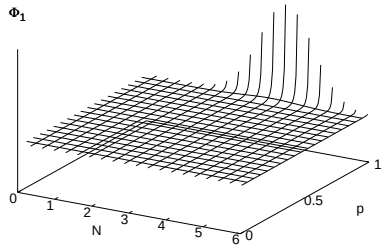
The merit of the marginal stationary distributions of N, $\phi_{N,1}$ and p, $\phi_{p,1}$, is of course that they are easier to compare with empirical data than the joint distribution Ψ_1. The marginal distributions $\phi_{N,1}$, and $\phi_{p,1}$ can be found from Ψ_1 by integrating out the unwanted variable. For instance, to find $\phi_{N,1}$ we solve the



(a) *Drift*



(b) *Migration*



(c) *Selection*

Figure 4.1 *Single-locus (unnormalised) stationary distribution* Φ_1 *with* $\hat{p} = 0.5$*,* $\eta = 1$*. (a) High drift with* $M = 0.001$*,* $s = 0.00001$*, (b) high migration with* $M = 3$*,* $s = 0.00001$ *and (c) strong selection with* $s = 3$*,* $M = 0.01$*.*

following integral,

$$\phi_{N,1} \propto N^{2M\zeta-1} e^{2N\zeta(1-s\eta)-\zeta N^2} \int_0^1 e^{4\zeta Nsp} p^{4M\zeta\eta\hat{p}-1} q^{4M\zeta\eta\hat{q}-1} \, \mathrm{d}p. \tag{4.9}$$

Using the definition of a confluent hypergeometric function (of the first kind,

Abramowitz & Stegun 1972),

$$ {}_1F_1\Big(a,\, b;\, c\Big) = \frac{\Gamma\Big(b\Big)}{\Gamma\Big(b-a\Big)\Gamma\Big(a\Big)} \int_0^1 e^{cp} p^{a-1} (1-p)^{b-a-1}\, \mathrm{d}p, \tag{4.10} $$

where $\Gamma\Big(a\Big)$ is the Gamma function (Abramowitz & Stegun 1972), and noting that $\overline{r}_{min} = (1 - s\eta)$ for one locus, the exact expression (up to a normalisation constant) for ϕ_N is

$$ \phi_{N,1} \propto N^{2M\zeta-1} e^{-\zeta(N-\overline{r}_{min})^2}\, {}_1F_1\Big(4M\eta\zeta\hat{p},\, 4M\eta\zeta;\, 4\eta\zeta N s\Big). \tag{4.11} $$

To find ϕ_p, the integral we solve is

$$ \phi_{p,1} \propto p^{4M\zeta\eta\hat{p}-1} q^{4M\zeta\eta\hat{q}-1} \int_0^\infty N^{2M\zeta-1} e^{2N\overline{r}-\zeta N^2}\, \mathrm{d}N. \tag{4.12} $$

Using integration by parts to solve the integral, the final result is

$$ \phi_p \propto p^{4M\zeta\eta\hat{p}-1} q^{4M\eta\zeta\hat{q}-1} Q(\zeta\overline{r}, \zeta M), \tag{4.13} $$

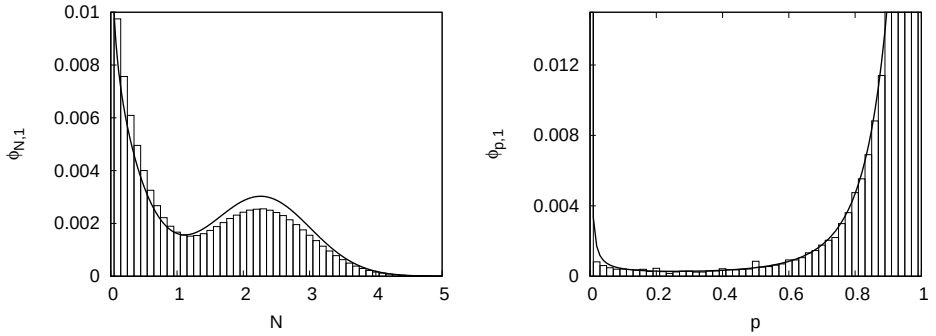
where function $Q(a,\, b)$ is defined as

$$ \begin{aligned} Q(a,\, b) = \frac{1}{2}\Big\{ & 2a\Gamma(\tfrac{1}{2}+b)\, {}_1F_1\Big(\frac{1}{2}+b,\, \frac{3}{2};\, a^2\Big) \\ & + \Gamma(b)\Big[{}_1F_1\Big(1+b,\, \frac{1}{2};\, a^2\Big) - 2a^2\, {}_1F_1\Big(\frac{1}{2}+b,\, \frac{3}{2};\, a^2\Big)\Big]\Big\}, \end{aligned} \tag{4.14} $$

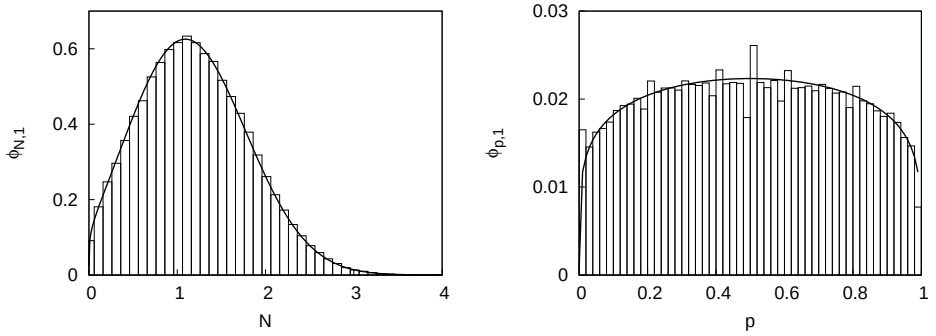
Figure 4.1 shows the unnormalised stationary distribution Φ_1 when random drift, migration and selection separately dominate the dynamics.

High random drift, $\zeta M,\, \zeta s \ll 1$

Figure 4.1a shows Φ_1 when drift dominates selection and migration. High drift causes frequent extinctions; it diminishes the relative effect of selection and migration (see Ch. 2). Furthermore, density regulation, produced by the $e^{-\zeta N^2}$ term in Eq. 4.7b, becomes severe. This explains the accumulation of probability mass at the extinction boundary $N = 0$. Drift also removes genetic variation from the population and the allele frequency distribution assumes the typical U-shaped form. Since selection is effectively absent under high drift, the local allele is likely to be lost due to extinction before selection has a chance to drive it to high frequencies. A high extinction rate also increases the probability mass at $p = 0$, a feature which is absent from the classical Wright-Fisher distribution



(a) $s = 3$, $M = 0.4$, $\hat{p} = 0.01$ **(b)** $s = 3$, $M = 0.4$, $\hat{p} = 0.01$



(c) $s = 0$, $M = 0.6$, $\hat{p} = 0.5$ **(d)** $s = 0$, $M = 0.6$, $\hat{p} = 0.5$

Figure 4.2 *Single-locus marginal distributions $\phi_{N,1}$ (left) and $\phi_{p,1}$ (right) for $\zeta = 1$ (diploids). Top, $\hat{p} = 0.01$, $s = 3$ and $M = 0.4$. Bottom, $\hat{p} = 0.5$, $s = 0$ and $M = 0.6$. The bars indicate simulation data ($K = 100$, $r_0 = 1/K$).*

(Wright 1931, 1932).

High migration, $\zeta M \gg 1$

High migration has two important effects. First, its demographic effect is to boost population size, which pulls the distribution away from the extinction boundary. Second, migration pulls the allele frequency distribution towards that of the migrant pool, homogenizing the metapopulation. If the local allele is almost fixed or lost on the mainland, Φ_1 becomes highly skewed towards either the $p = 1$ or $p = 0$ boundary respectively. Otherwise Ψ_1 has a single bivariate, Gaussian-type peak. This is shown in Fig. 4.1b. The power of migration is clearly illustrated by its effect on $\phi_{N,1}$ in the absence of selection, see Fig. 4.2c. Provided the migration rate satisfies $M > \frac{1}{2}$, migration need not be relatively high for the probability of extinction to be significantly reduced. Large migration rates are needed, however, to pull the allele frequency distribution towards the migrant average, $\hat{p}$, see Fig. 4.2d.

Strong selection, $\zeta s \gg 1$

Figure 4.1c shows Φ_1 when selection dominates. Strong selection drives the population towards maximum mean fitness, hence high allele frequencies. Even when drift is stronger than selection and the local allele is rare, the effect of selection is felt through $\phi_{p,1}$, which becomes skewed towards $p = 1$, see Fig. 4.2b. The demographic effect of selection is tied up with growth: a higher mean fitness results in larger population sizes. However, strong selection cannot save the population from low numbers, or indeed extinction, if migration is too low, see Fig. 4.2a.

To summarise, migration and low random drift are key for the demic population's existence. The fate of the favoured allele is primarily dependent on the strength of these forces, since without persistence there is no population within which the favoured allele can be established. Selection counteracts drift and maladapted gene flow by increasing mean fitness and hence population size. Strong selection boosts the chances of survival, making it more likely the local allele will spread. We next investigate adaptation when the allele is rare on the mainland, $\hat{p} \sim 0$. We make the important assumption of weak random drift. By 'weak' we mean ζs, $\zeta M \gg 1$. There are two reasons for this. First, weak random drift is synonymous with 'small' or 'weak noise'. In the limit of weak noise the stochastic behaviour of the population closely follows its deterministic behaviour (Van Kampen 2007), greatly simplifying the analysis. Second, when random drift is high, a diffusion description cannot adequately approximate the corresponding discrete models: the diffusion process is accurate only for small generational changes, whereas discrete generational changes are likely to be large for high drift.

4.3.2 Rare allele case

In the last chapter it was found that when the local allele is rare on the mainland, for a single, diploid locus under high migration, strong selection relative to growth ($s > 1$), there are two possible deterministic equilibria: a polymorphic state with a large population size (high state), due to migration-selection-death balance, and a maladapted state with a smaller population size (low state), due to the swamping effect of migration, see §3.3.4. For an intermediate range of migration rates both states are bistable. If the favoured allele is completely lost ($\hat{p} = 0$), this range is defined exactly by the critical rates M_l and M_h.

The stochastic equivalent of these results is bimodality in the stationary

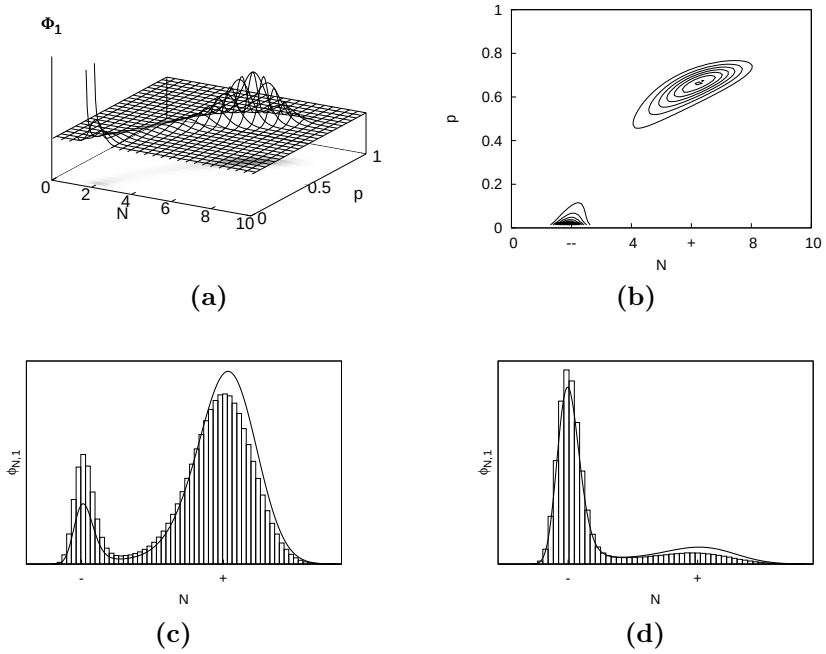


Figure 4.3 *Bimodality in in the single-locus stationary distributions. (a) the (unnormalised) stationary distribution* Φ_1 *for parameters* $\hat{p} = 0.002$*,* $s = 8$*,* $M = 17$*,* $\zeta = 1$ *(diploid).(b) a contour plot of* Φ_1*. Bottom row, the marginal distributions of* ϕ_N *for (c)* $M = 16$*. and (d)* $M = 18$ *(right). The bars indicate simulation data (*$K = 500$*).*

distribution Ψ_1. For $\zeta s \gg 1$ and $\zeta M \gg 1$, Ψ_1 clusters about a low allele frequency, low population density 'maladaptive peak' ('-') and a high allele frequency, large population density 'well-adapted' peak ('+'). An example of bimodality is shown in Fig. 4.3a. Because $\hat{p}$ is low, the maladapted 'peak' actually appears as a divergence at boundary $p = 0$. We will regard such a divergence as a 'boundary peak'. In contrast, the well-adapted peak which results from migration-selection-drift balance is an 'interior peak', far from any boundaries.

When fluctuations are small, we expect the stationary results to closely parallel the deterministic behaviour of the population. For a given selection strength and for migration rates that approximately lie within the deterministic bistable range, Ψ_1 should be bimodal. Each deterministic attractor delineates a different 'ecological' (or 'adaptive') zone (see the contour plot, Fig. 4.3b). Within the domain of the high state attractor, the population is within its ecological niche (Hutchinson 1958). We define 'niche' (Hutchinson 1958) here to mean a habitat in which the population experiences a positive growth rate on average together with other conditions that are conducive to its survival; for example a low extinction probability and high allele frequency. This is the 'well-adapted'

ecological zone. Within the domain of the low state attractor however, the population is vulnerable to extinction due to low population numbers and is poorly adapted with low mean fitness. This is the 'maladaptive' ecological zone. Fluctuations can send the mean fitness below zero whilst the population occupies the maladaptive zone; thus the maladaptive zone is outside the species' niche (Holt 2009).

The location of the peak maxima should be close to its corresponding deterministic state. To be more precise, the extrema of Ψ_1 are,

$$N^*_{\pm} = \frac{\overline{r} \pm \sqrt{(\overline{r})^2 + 2(2M - \frac{1}{\zeta})}}{2} \tag{4.15a}$$

$$p^*_{\pm} = \frac{(Ns - M + \frac{1}{2\zeta\eta}) \pm \sqrt{(Ns - M + \frac{1}{2\zeta\eta})^2 - 2Ns(2M\hat{p} - \frac{1}{2\zeta\eta})}}{2Ns}. \tag{4.15b}$$

When ζs and ζM are large, fluctuations are small, and from Eq. 4.15 we can see that the peak maxima do indeed coincide with the deterministic equilibria for low drift, cf. Eq. 3.1. However, bimodality also exists when random drift is moderately strong, ζs, $\zeta M \approx 1$. If selection and migration are strong, ζs, $\zeta M \gg 1$, bimodality can also exist for haploids; recall that in the deterministic case, the haploid equilibrium is monostable. In this case, the correspondence with the deterministic states from the previous chapter disappears. This is because boundary peaks form when $\hat{p}$ is small or M is small (recall that there are no absorbing states when $\hat{p} > 0$ and $M > 0$). We refer to such boundary peaks as 'ghost peaks'. A 'ghost peak' does not have a deterministic state associated with it, since stochasticity can shift the peak maximum beyond biologically valid limits. For example, p^* may be negative, which clearly has no valid deterministic counterpart. In such a case, the maxima of the peak can be crudely approximated as the closest corresponding deterministic state when fluctuations are small.

The relative degree of stability of the two peaks is defined by their relative densities (Gardiner 1985). This changes over the range of migration rates for which Ψ_1 is bimodal. As shown in figures 4.3c and 4.3d, the peak densities of marginal distribution $\phi_{N,1}$ varies considerably over a small change in rate M. As migration varies, the domain of one attractor grows as the domain of the alternate attractor shrinks. Similarly, the stochastic peak densities grow and shrink with varying M. Therefore the probability of occupying each zone will also increase and diminish with rate M. We consider an ecological zone to 'exist' if there is a non-negligible probability of it being occupied by the population.

We proceed by looking at the stationary probability that a particular zone will be occupied by the population, i.e. the 'peak probability'. We then calculate the stationary moments of N and p. We will use them to characterise the 'bimodal range', defined as the range of migration rates for which both ecological zones exist. With these stationary quantities we can find the conditions necessary for local adaptation.

Exact calculations using distribution Ψ_1 are intractable. However, by assuming that random drift is low, it is possible to approximate Ψ_1 using the local peak approximations and this is what we discuss next.

Local peak approximations

The basic premise is as follows: any general multi-peaked, multivariate distribution described by a potential can also be approximated as a sum of distinct peaks, provided the peaks are narrow and well-separated. Each peak k can then be approximated by a local Taylor expansion of the potential about the peak's maximum $\mathbf{x}_k$. For central or interior peaks, such as the adaptive peak, this is usually equivalent to a saddle-point or local Gaussian approximation (Bender & Orszag 1991). Boundary peaks, being distinctly non-Gaussian, require delicate treatment. Details can be found in App. B, §B.0.3. This approximation scheme has been applied to one-dimensional (Lande 1985*a*) and multidimensional genetic distributions (Barton & Rouhani 1987*a*).

We set the maximum of each peak to be located at the corresponding deterministic equilibrium, which is a reasonable approximation in the limit of small fluctuations. Furthermore, we divide peaks along the deterministic unstable equilibrium between the peaks. These approximations are accurate to leading order in ζ^{-1}, but do surprisingly well even for $\zeta = 1$.

The generic form of the approximation of Ψ_1 is

$$\tilde{\Psi}_1 \approx \frac{1}{\sum_{j\in+,-} w_j \Phi_{1,j}} \sum_{k\in+,-} \Phi_{1,k} \tilde{\Psi}_{1,k} \tag{4.16}$$

where $\Phi_{1,k} = \Phi_1(\mathbf{x}_k)$ and w_k is the 'weight' of the k^{th} peak, $w_k = \int_{\sim \mathbf{x}_k} \Phi_1 \, \mathrm{d}\mathbf{x}$. Function $\tilde{\Psi}_{1,k}$ is the approximate (unweighted) form of peak k and depends on the Taylor expansion about the local maximum.

We try three different approximations. For two of them, the high state peak is approximated as a bivariate Gaussian. The low state peak requires more care and we try: 1) a bivariate Gaussian (the 'Gaussian' approximation) and 2) the product of two Gamma distributions, since the coupling between N and p is

s should be weak when $M \gg Ns$ (the 'Gamma' approximation). The use of Gamma distributions to approximate multidimensional boundary peaks was first discussed in Barton & Rouhani (1987*a*) for a multilocus model with constant population size. Because the spread of each peak is negligibly small relative to the distance between the two maxima in the limit of low drift, the third approximation is to treat each peak as a delta function, $\delta(\mathbf{x} - \mathbf{x}_k)$ (the 'zero-width' approximation).

Explicitly, the three approximations are

$$\tilde{\Psi}_G \approx \frac{1}{\Phi_{1,+}D_+ + \Phi_{1,-}D_-} \sum_{k\in+,-} \Phi_{1,k} e^{-\frac{1}{2}(\delta_k\mathbf{x})^T \mathbf{\Omega_k} (\delta_k\mathbf{x})}, \tag{4.17a}$$

$$\tilde{\Psi}_\gamma \approx \frac{1}{D_+\Phi_{1,+} + C_-\Phi_{1,-}} \Big(\Phi_{1,+} e^{-\frac{1}{2}(\delta_+\mathbf{x})^T \mathbf{\Omega}_k (\delta_+\mathbf{x})} + \Phi_{1,-} g_N(N) g_p(p) \Big), \tag{4.17b}$$

$$\tilde{\Psi}_0 \approx \frac{1}{\Phi_{1,+} + \Phi_{1,-}} \Big(\Phi_{1,+}\delta(\mathbf{x} - \mathbf{x}_+) + \Phi_{1,-}\delta(\mathbf{x} - \mathbf{x}_-) \Big) \tag{4.17c}$$

where $\delta_k\mathbf{x} = \mathbf{x} - \mathbf{x}_k$ and

$$\Omega_{ij} = -2\partial_{x_j}\partial_{x_i} U \tag{4.18}$$

$$D_k = \sqrt{\frac{(2\pi)^n}{\det(\zeta\Omega)|_{\mathbf{x}=\mathbf{x}_k}}} \tag{4.19}$$

$$g(x;a,\,b) = e^{-a(x-x_-)} \Big(\frac{x}{x_-} \Big)^b \tag{4.20}$$

$$g_N(N) = g(N; 2\zeta|\overline{r}_{min}|,\, 2M\zeta) \tag{4.21}$$

$$g_p(p) = g(p; 4M\zeta\eta,\, 4M\zeta\eta\hat{p}) \tag{4.22}$$

$$C_- = \frac{1}{\Psi_-} h(2\zeta|\overline{r}_{min}|,\, 2M\zeta) h(4M\zeta\eta,\, 4M\zeta\eta\hat{p}) \tag{4.23}$$

$$h(a,\,b) = a^{-b}\Gamma(b). \tag{4.24}$$

Note that x_+ (x_-) is the value of x at the maximum of the '+' ('-') peak. Also, $w_k = 1\ \forall\ k$ in the zero-width approximation. Note that matrix Ω_{ij} is related to the Jacobian of the deterministic dynamics (see §B.0.3).

Next we consider the likelihood of finding the population within a particular ecological zone at equilibrium. Numerical results of the exact probabilities are compared with approximate expressions derived from each of the three different peak approximations. We then move on to calculate the stationary moments of N and p.

Probability of occupying an ecological zone

What is the stationary probability that the population is well-adapted? This 'zone occupation' probability is the same as the probability of finding the population near the well-adapted peak. Assuming peaks are well-separated, for peak k, this is just $P(k) = \int_{\sim \mathbf{x}_k} \Psi_1 \, \mathrm{d}\mathbf{x} = w_k / Z_0$. The approximate peak probabilities have the generic form

$$P(+) = \left(1 + \frac{w_-}{w_+} \frac{\Phi_{1,-}}{\Phi_{1,+}}\right)^{-1}, \tag{4.25a}$$

$$P(-) = \left(1 + \frac{w_+}{w_-} \frac{\Phi_{1,+}}{\Phi_{1,-}}\right)^{-1}. \tag{4.25b}$$

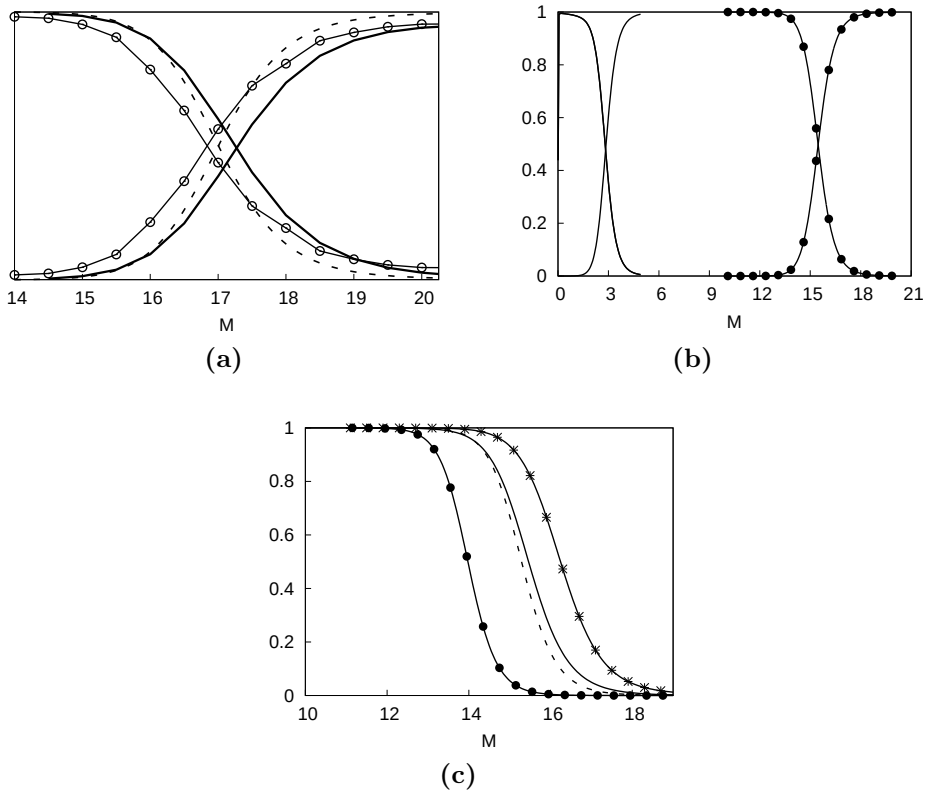


Figure 4.4 *Single-locus zone probabilities $P(+)$ (decreasing curves) and $P(-)$ (increasing curves) for varying migration rate M. a) $s = 8$ and $\hat{p} = 0.002$ with the zero-width approximation (dashed lines), exact (continuous) results (solid lines), and simulation results (with open circles) for $K = 500$. Note the error bars are too small to be visible. b) exact results for $\hat{p} = 10^{-4}$ and $s = 4$ (solid lines), and $s = 8$ (with circles). c) $P(+)$ for $s = 8$ with exact results (solid lines) and zero-width (dashed lines), Gaussian (with crosses) and Gamma (with filled circles) approximations.*

The exact results, approximations and simulation results are shown in Fig. 4.4. By 'exact' we mean exact in the continuous model only. The exact probabilities were calculated by numerical integration and all numerical computations were carried out using MATHEMATICA (Wolfram 1999). The distribution was separated into two peaks along the unstable deterministic equilibrium. The results were insensitive to the small differences between the deterministic unstable equilibria over the range of migration rates shown.

From Fig. 4.4a we see that the probability $P(+)$ of adaptedness decreases steadily for increasing migration rates, until it rapidly decays over a small range of rates, the 'bimodal range', as maladaptation becomes more probable. The maladaptive zone begins to make an appearance as Ψ_1 becomes bimodal. The transition to bimodality is not sharp, in contrast to the abrupt onset of bistability in the deterministic model. The population is equally likely to be found in a well-adapted or maladapted state at a critical migration rate M_{max}. The migration rate just below M_{max} is the largest rate at which the population can still experience a bias towards being found within its niche. Above M_{max} the population's chances of being found within its ecological niche disappear and Ψ_1 is singly-peaked. Note that despite the certainty of maladaptation and sink-status, the stationary probability of extinction is population is zero, see Fig. 4.3d: migration saves the population from extinction, the "rescue effect" (Hanski 1999).

In Fig. 4.4b the peak probabilities are shown for different selective strengths s. The stronger selection is, the higher the value of rate M_{max} and the stronger migration has to be for the population to be in a state of maladaptation.

The zero-width approximation is shown in Fig. 4.4a. Agreement between the exact result and the zero-width approximation does reasonably well considering the shape of the peaks are not taken into account. In Fig. 4.4c, the exact result and results from all three approximation schemes for $P(+)$ are shown. The Gaussian approximation does less well, but this is expected since the low peak is not locally Gaussian. For decreasing fluctuations these two approximations appear to converge. This is because the width of the peaks become narrower, relative to the peak heights and the distance between the peaks. The Gamma approximation does poorly. This is because the low peak cannot be reduced to a product of two Gamma distributions. (Note that the Gamma approximation worked in Barton & Rouhani (1987*a*) precisely because there was no coupling between the allele frequencies in their constant population size model).

Finally, the discrete results from MWFM are in reasonable agreement with theory; the systematic 'error' of both curves being shifted to the left is due to

an inherent difference between the discrete and continuous models, see §4.4.4. Essentially, in the discrete model, selection is less robust against migration and occupation of the maladaptive zone becomes likely for lower rates than predicted by the continuous theory.

Stationary moments

We wish to find the stationary average, $\langle x_i \rangle = \int x_i \Psi_1 \, \mathrm{d}\mathbf{x}$, and variance, $\sigma^2_{x_i} = \int (x_i - \langle x_i \rangle)^2 \Psi_1 \, \mathrm{d}\mathbf{x}$ of N and p. The stationary moments are important because, for example, they tell us when the population experiences the greatest genetic or demographic variation. Moreover, they are, in principle at least, empirically measurable quantities. They also are good indicators of bimodality, as we shall see.

Exact analytical expressions for the moments can only be found for the neutral case ($s = 0$), provided $\hat{p} > 0$. The latter condition prevents divergence of the integrals over N and p. Using the marginal distributions in Eq. 4.11 and Eq. 4.13, the first two moments are

$$
\begin{aligned}
\langle N \rangle &= C_N \frac{1}{2} \zeta^{-\zeta M} \Big[\zeta^{-\frac{1}{2}} \Gamma\Big(\frac{1}{2} + \zeta M\Big) {}_1F_1\Big(\frac{1}{2} + \zeta M,\, \frac{1}{2};\, \zeta\Big) \\
&\quad + 2\Gamma\Big(1 + \zeta M\Big) {}_1F_1\Big(1 + \zeta M,\, \frac{3}{2};\, \zeta\Big) \Big] \\
\langle p \rangle &= \hat{p} \\
\sigma^2_N &= C_N \Big[\frac{1}{2} \Gamma\Big(1 + M\Big) {}_1F_1\Big(1 + M, \frac{1}{2}, 1\Big) + \Gamma\Big(\frac{3}{2} + M\Big) {}_1F_1\Big(\frac{3}{2} + M, \frac{3}{2}, 1\Big) \Big] \\
\sigma^2_p &= \frac{\Gamma\Big(4\zeta\eta M\Big) \Gamma\Big(2 + 4\zeta\eta M\hat{p}\Big)}{\Gamma\Big(4\zeta\eta M\hat{p}\Big) \Gamma\Big(2 + 4\zeta\eta M\Big)},
\end{aligned}
$$

where

$$
C_N = 2\zeta^{\zeta M} \Big(\Gamma\Big(\zeta M\Big) {}_1F_1\Big(\zeta M,\, \frac{1}{2}, \zeta\Big) + 2\sqrt{(}\zeta) \Gamma\Big(\frac{1}{2} + \zeta M\Big) {}_1F_1\Big(\frac{1}{2} + \zeta M, \frac{3}{2}, \zeta\Big) \Big)^{-1},
$$

Expressions for higher moments can also be found exactly, but these expressions quickly become unwieldy.

Adaptation cannot happen without selection. However, it is still instructive to inspect the neutral case as a basis of comparison once selection is taken into account. It also allows us to isolate the effect of migration. With selection, approximate expressions for the stationary moments can be found using the

generic form of the approximation $\tilde{\Psi}_1$ in Eq. 4.16. These are,

$$\langle x \rangle = \frac{1}{\sum_j w_j \Phi_{1,j}} \Big(\sum_k w_k \Phi_{1,k} x_{i,k} \Big), \tag{4.26a}$$

$$\sigma^2_{x_i} = \frac{1}{\sum_j w_j \Phi_{1,j}} \Big(\sum_k w_k \Phi_{1,k} (\sigma^2_{x_{i,k}} + x^2_{i,k}) - \langle x_i \rangle^2 \Big). \tag{4.26b}$$

where $\sigma^2_{x_k}$ can be found from matrix $\mathbf{\Omega}$ evaluated at $\mathbf{x}_k$ (see App. B for details). Explicit expressions for the first two raw moments of N and p are

$$\langle N \rangle = \triangle_\cap N + \frac{1}{2} \triangle_\cup N \tanh\Big(2\triangle_\cup U + \ln\Big(\frac{w_+ \Phi_{1,+}}{w_- \Phi_{1,-}} \Big) \Big), \tag{4.27}$$

$$\langle p \rangle = \triangle_\cap p + \frac{1}{2} \triangle_\cup p \tanh\Big(2\triangle_\cup U + \ln\Big(\frac{w_+ \Phi_{1,+}}{w_- \Phi_{1,-}} \Big) \Big), \tag{4.28}$$

and

$$\langle N^2 \rangle = \triangle_\cap Y_N + \frac{1}{2} \triangle_\cup Y_N \tanh\Big(2\triangle_\cup U + \ln\Big(\frac{w_+ \Phi_{1,+}}{w_- \Psi_{1,-}} \Big) \Big), \tag{4.29}$$

$$\langle p^2 \rangle = \triangle_\cap Y_p + \frac{1}{2} \triangle_\cup Y_N \tanh\Big(2\triangle_\cup U + \ln\Big(\frac{w_+ \Phi_{1,+}}{w_- \Psi_{1,-}} \Big) \Big), \tag{4.30}$$

where $\triangle_\cap f = \frac{1}{2}(f_+ + f_-)$, $\triangle_\cup = (f_+ - f_-)$ and $Y_{x_i} = \sigma^2_{x_i} + x^2_i$, where f_k denotes evaluation of f at $\mathbf{x}_k$. From these moments the variances can of course be calculated.

Figure 4.5 shows the stationary moments for the neutral case when the favoured allele is rare ($\hat{p} = 0.01$). The average population size increases with increasing migration, and the average allele frequency remains approximately constant at $\hat{p}$, see figures 4.5a and 4.5c. Demographic and genetic variation steadily decrease with increasing migration, see figures 4.5b and 4.5d. Simulation data from the MWFM are also shown and agreement with the continuous model is good for the averages. On the other hand, the variances are in good agreement with theory only for large carrying capacities, for example $K = 100$ compared with $K = 20$ in figures 4.5b and 4.5d. The reasons for this will be discussed in §4.4.4.

In Wright's island model, with constant effective population size N_e and gene flow rate m, the expected allele frequency is $\langle p \rangle_W = \hat{p}$ and the variance is $\sigma^2_{p,W} = \frac{\hat{p}(1-\hat{p})}{4N_e m + 1}$ (Wright 1931, 1932; see Ch. 2). Using $M = N_e m$, Wright's results are

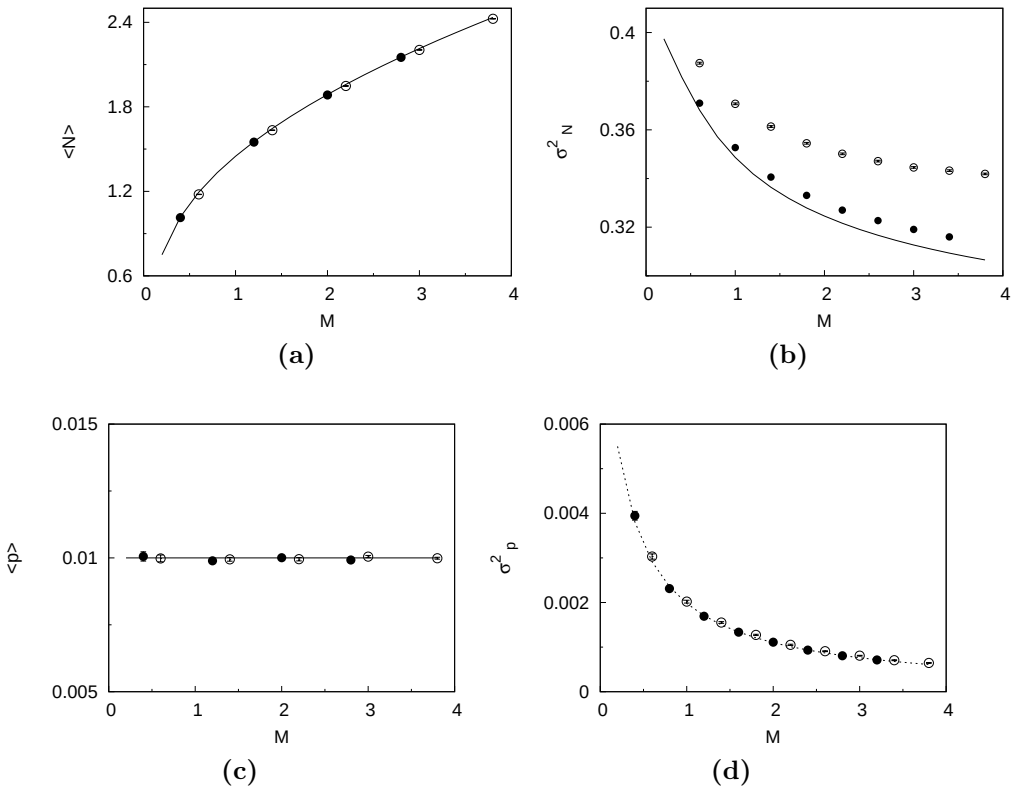


Figure 4.5 *Neutral stationary moments for one locus for N (top row) and p (bottom) with $\hat{p} = 0.01$, $\zeta = 1$. Exact results (solid line) are compared with simulation data from the MWFM (circles) for $K = 100$ (filled symbols) and $K = 20$ (open symbols) with $r_0 = \frac{1}{K}$. In (d) theory (dotted line) exactly matches $\sigma^2_{p,W}$ (Wright's variance for constant N, dashed line).*

compared with our theory in figures 4.5c and 4.5d. The perfect match between the moments is not surprising when we recall that N does not feature at all in the marginal allele frequency distribution, $\phi_{p,1}$ Eq. 4.13, and is identical to Wright's distribution (if $M = N_e m$).

The stationary moments under selection are markedly different from the neutral case, as shown in Fig. 4.6. Over a small range of migration rates, the average population size and allele frequency smoothly vary from large values, corresponding to probable occupation of the ecological niche, to low values, corresponding to maladaptation (see figures 4.6a and 4.6c).

These curves are similar to the high and low deterministic equilibria, also included in the figures. Fluctuations smooth out the transition between an average state of adaptedness and an average state of maladaptedness, in contrast to the corresponding transition in the deterministic case. The population is most vulnerable when the average population size is at its lowest; this is when the

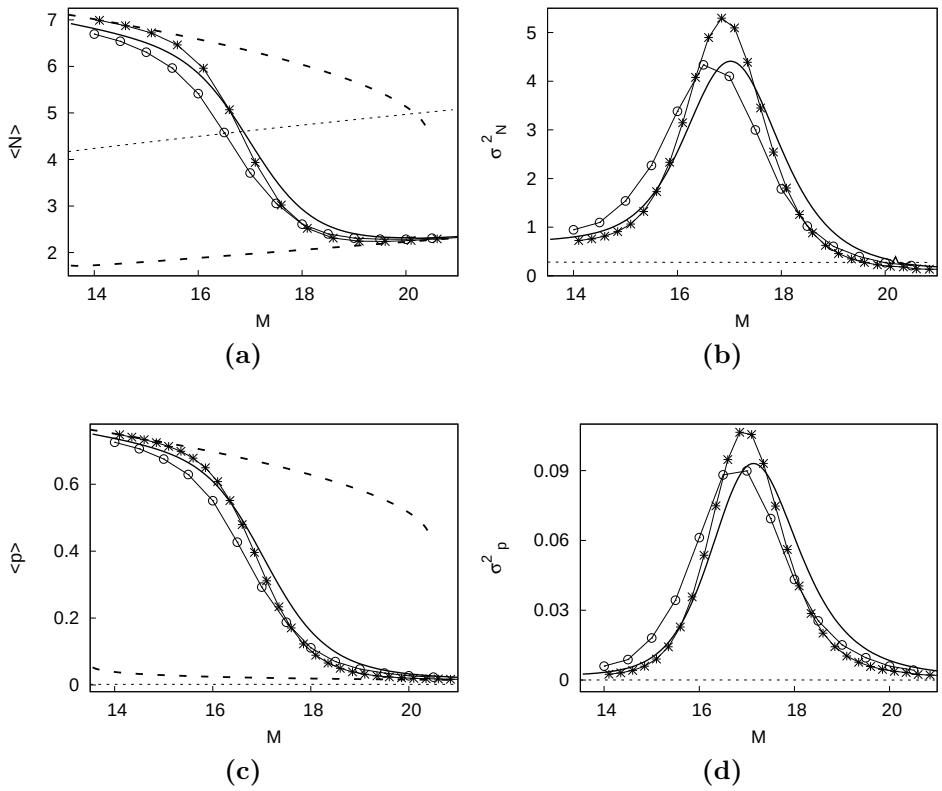


Figure 4.6 *Single-locus stationary moments for $s = 8$, $\hat{p} = 0.002$: N (top row) and p (bottom row) with $\zeta = 1$. Numerical results (solid lines), approximate results from $\tilde{\Psi}_0$ (with crosses) and simulation data (with open circles) for $K = 500$. Deterministic equilibria (long-dashed lines) and the neutral moments (short-dashed lines) are also shown.*

ecological niche only just disappears and before migration raises the average population size, see Fig. 4.7a.

Comparing the genetic moments with their neutral counterparts, the population fares better under selection, even though gene flow is maladaptive, for migration rates at the lower end of the bimodal range. The average population size however, falls below the neutral population size once the maladapted zone is more likely to be occupied. This is a direct consequence of hard selection.

Both variances peak at approximately rate $M = M_{max}$, see figures 4.7b and 4.7d; local adaptation is possible when migration rates are either below or within this range. Genetic and demographic variation is greatest in the population when there is no bias towards occupation of either ecological zone (i.e. the zone probabilities are $P(+) = P(-) = \frac{1}{2}$). An abundance of genetic variation means the population can in principle adapt to unexpected changes in its environment or circumstances. Therefore the population has its greatest chance of adapting

in the face of maladaptive gene flow and drift at migration rate M_{max}.

Agreement between theory, the discrete model and approximations using the zero-width scheme is good. For migration rates much below or above M_{max}, the averages are indistinguishable from the deterministic equilibria. It is not so surprising that the deterministic equilibria and the approximations based on them are in good agreement; for $\zeta s = 8$ and $\zeta M > 14$, selection and migration are strong and fluctuations are significantly damped.

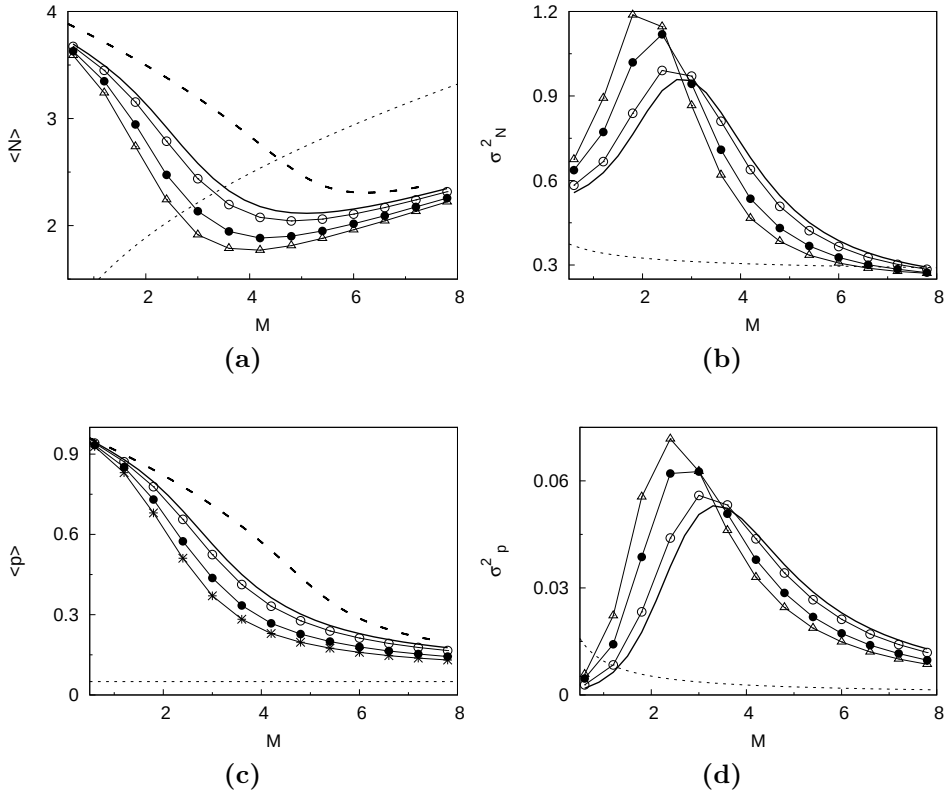


Figure 4.7 *Single-locus stationary moments for $s = 3$, $\hat{p} = 0.05$ and $\zeta = 1$. Numerical results (solid lines) and simulation data from the MWFM for $K = 100$ (with open circles), $K = 30$ (with filled circles) and $K = 20$ (with open triangles) are shown. Deterministic equilibria (long-dashed lines) and the neutral moments (short-dashed lines) are also shown.*

Contrast these stationary moments with moments for weaker selection and migration, shown in Fig. 4.7. The population is obviously more vulnerable to maladaptation and at low migration rates ($\zeta s > \zeta M$), the averages rapidly decline. We cannot approximate the stationary moments in this case because the deterministic equilibria are significantly different from the maxima of Ψ_1; in the deterministic model, bistability does not exist for these parameter values. Simulation data from the MFWM agrees well with theory but only for carrying

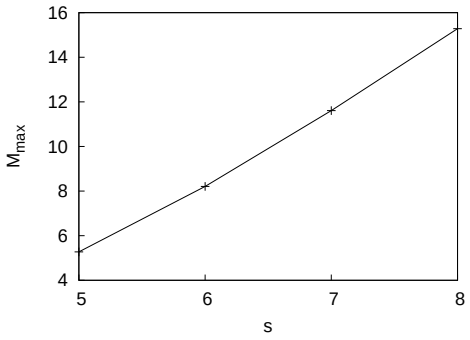


Figure 4.8 *Migration rate M_{max} for one locus when zone occupation is equiprobable and the variances of N and p are at a maximum.*

capacity $K = 100$. This seems to be the carrying capacity at which the discrete nature of the simulated population can be approximated as smoothly varying. Despite the quantitative discrepancies, data for smaller carrying capacities still capture the qualitative features of the stationary moments.

We next use the peak probabilities and the stationary moments to define the 'bimodal range', that is, the range of migration rates for which occupation of either ecological zones is non-negligible.

Defining the bimodal range

The bimodal nature of Ψ_1 is appreciable only for a narrow range of migration rates. In this narrow range, the averages deviate markedly from the deterministic equilibria and the variances are large. This occurs for migration rates close to M_{max}, the migration rate for which the population is equally likely to be found in either zone.

In the last chapter, bistability existed for migration rates between the critical values M_l and M_h. The 'bistable width' of the range was defined as $\triangle M = M_u - M_h$ (see §3.3.4). A similar range is applicable to the stochastic case when drift is low. To leading order in $\frac{1}{\zeta}$, within this range, occupation of either ecological zone is possible. However, for much of this range, the occupation probability of one zone is much greater than the other zone. Therefore the bimodal width, is much smaller than the corresponding bistable width, see for example Fig. 4.6a. We could define the bimodal range to be $\triangle_\zeta M = M_{max} \pm \delta M$, where $M_{max} \pm \delta M$ can be arbitrarily defined as the rates at which the variances are at full-width-half-maximum values, or when one of the two peak probabilities falls below a cutoff.

To find M_{max} exactly we would need to invert the ratio of peak densities when

the peak densities are equal; if $R = \Psi_{1,+}/\Psi_{1,-}$, we would need to invert $R = 1$. This can only be done numerically and is rather cumbersome. Alternatively, we can use the peak probabilities to find M_{max} by fitting the functions $g_{\pm}(x) = \frac{1}{2}[1 \pm \tanh(x - M_{try})]$ to the probability curves $P(+)$, $P(-)$ and determining the ordinate M_{try}, where the curves intersect, i.e. where $g_{\pm}(M_{try}) = P(+) = P(-) = \frac{1}{2}$. In Figure 4.8 the results from this fitting is shown. Rate M_{max} has the expected trend of increasing for strong selection.

We have considered the case when the local allele is rare on the mainland, $\hat{p} \sim 0$. We end discussion of the single-locus stationary properties by considering migration when migrants are polymorphic and the local allele is more frequent on the mainland, $0 \ll \hat{p} < \frac{1}{2}$.

4.3.3 Polymorphic gene flow

As shown in Fig. 4.9, when the local allele is not rare on the mainland, though still not common, $\hat{p} \ll \frac{1}{2}$, bimodality disappears and so does the distinction between the two ecological zones. There is a gradual degradation of the demic state from a highly fit population on average, to a state of maladaptation, see Fig. 4.9c. This is unlike the smooth but rapid transition between the well-adapted and maladapted zones when the local allele is rare. The demographic average increases, however, and the population is saved from extinction at low numbers, see Fig. 4.9a. Genetic variation also increases with high migration, which again has implications for future adaptability, see Fig. 4.9d.

Once again, simulation results for $K = 100$ are in better agreement with numerical results than data for $K = 20$. However, the results for the smaller carrying capacity are in better agreement with theory than they are for the neutral moments with a lower migrant frequency, see Fig. 4.5. The stationary moments appear to be converging towards the deterministic equilibria together with the neutral moments. The exception is the average allele frequency, which requires very large migration rates to show appreciable convergence.

The stationary results show that under highly maladaptive gene flow, for migration rates within a narrow range of rates, the population can be found in one of two ecological zones. Across a range of rates, the bimodal range, the chance of finding the population within a particular zone varies considerably, as do the stationary moments of N and p.

We next consider the first-passage properties of the stochastictrajectories.

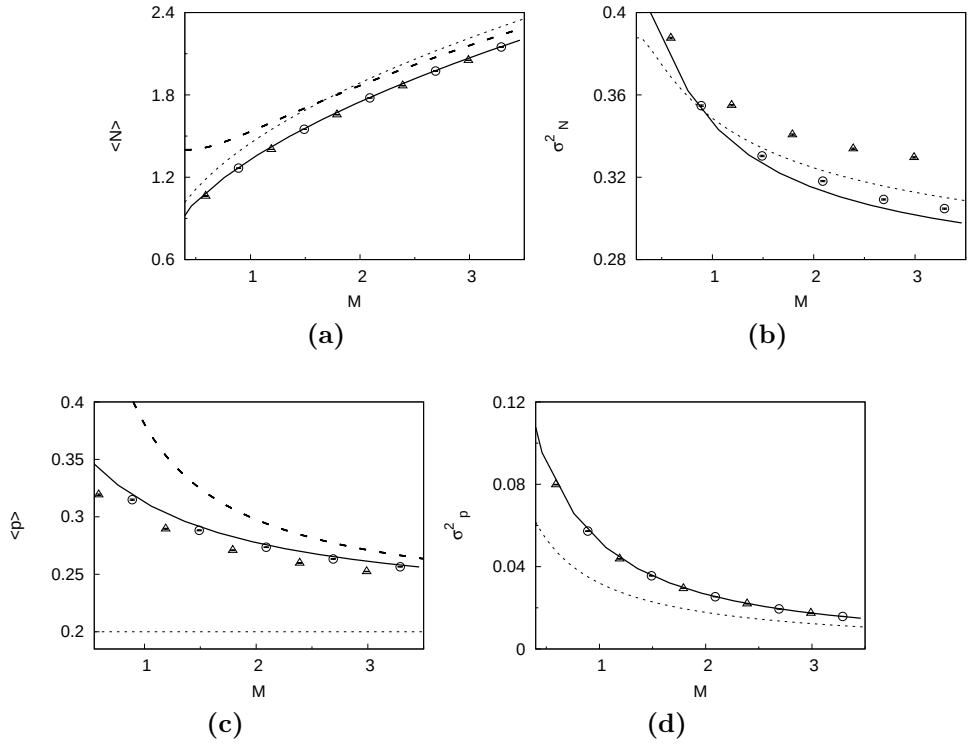


Figure 4.9 *Stationary moments for one locus with $s = 0.5$, $\hat{p} = 0.2$ and $\zeta = 1$. Numerical results (solid lines), approximate results from $\tilde{\Psi}_0$ (with crosses) and simulation data for $K = 100$ (open circles) and $K = 20$ (open triangles).*

4.3.4 Stochastic trajectories

The stationary properties of the previous section describe the characteristics of an ensemble of identically evolving demes under similar conditions, or equivalently the behaviour of a single population that persists indefinitely. In nature, however, there are no such ensembles and a population may not persist long enough to reach stochastic equilibrium. Therefore properties of the stochastic trajectories are important for understanding adaptation and survival. First passage properties characterise the first time an event occurs. In this section we consider the mean extinction time for the neutral and non-neutral case. We then consider the probability of fixation. These first passage properties are important when migration and selection are weak relative to drift. When selection and migration are strong, within the appropriate bimodal range of migration rates, the population is able to transition between the ecological niche and maladaptive zone. We attempt to calculate an asymptotic upper bound for the rate of transitions between zones.

Mean extinction time

All populations eventually become extinct. The important question is then not whether but when. In this section, we briefly investigate the mean time until extinction (MET), the average time it takes for extinction to occur (Redner 2001) when the population is isolated and subject to both demographic and genetic processes. Specifically, we look at continuous approximations of the mean extinction time in the discrete Modified Wright-Fisher model. We will reproduce the standard calculation of the MET for the neutral case, when only demographic processes cause extinction; this result is then compared to simulation data. We then use simulations to gain insight into the effect of selection on the MET. Clearly, being able to predict the mean time until a population falls extinct given its initial size and additive genetic variance, has potential applications in conservations studies.

The MET is a popular quantifier of extinction risk in population biology and conservation studies (Soulé & Simberloff 1986, Scahffer 1987, Lande 1988, Mangel & Tier 1993*b*, Lande et al. 2003) and is particularly important for small populations, which experience strong demographic stochasticity (Lande et al. 2003, Gabriel & Bürger 1992). Extensive work has been carried out on the MET in genetically monomorphic populations subject to solely demographic processes. For example, it has been calculated for single populations under density independent (MacArthur & Wilson 1967, Richter-Dyn & Goel 1972) and density dependent growth (Ludwig 1976, Tier & Hanson 1981, Leigh 1981), with sex ratios (Gabriel & Bürger 1992), Allee effects (Brassil 2001), environmental stochasticity and random catastrophes (Mangel & Tier 1993*a*, Lande 1993, Halley & Iwasa 1998). The effect of migration on the MET has been studied in unstructured populations (Mangel & Tier 1993*a*), as well as metapopulation models (Hill et al. 2002, Lande et al. 1998).

In comparison, little work has been done to explore extinction in an evolutionary context. The accumulation of deleterious mutations can lead to extinction through mutational meltdown (Lynch & Gabriel 1990). In a temporally varying environment, beyond a critical rate of long-term environmental change, extinction is certain for a population under stabilising selection and density dependent growth (Lynch et al. 1991, Lynch & Lande 1993, Bürger & Lynch 1995). The interplay between ecological and evolutionary processes can dramatically reduce the mean extinction time below values predicted by considering any single process in isolation (Bürger & Lynch 1995).

We investigate how the strength of selection and the initial genetic variation

in the population affects the average extinction time. But first we derive the neutral MET result.

For convenience we will revert to the *unscaled* variables (see §3.2.1). In the neutral case, the MET for a population of initial size N_0, $T_{met}(N_0)$, can be found exactly for the continuous model; the dynamics of N_u and p decouple and the problem reduces to one dimension. This has been found for a stochastic logistic model by Leigh (1981). The population size follows a diffusion process with infinitesimal mean and variance, $A(N_0) = r_0 N_0(1 - \frac{N_0}{K})$ and $B(N_0) = N_0$, (see Eq. 4.3 and Eq. 4.4 with $\zeta = 1$) defined on the interval $N_0 \in [0, \infty]$. $T_{met}(N_0)$ satisfies the inhomogeneous Backward Fokker-Planck equation (BFPE) (Karlin & Taylor 1975*b*)

$$-1 = A(N_0)\partial_{N_0} T_{met}(N_0) + \frac{1}{2} B(N_0)\partial^2_{N_0} T_{met}(N_0). \tag{4.31}$$

Two boundary conditions are needed to fully specify the diffusion. The extinction boundary can be defined to occur at a low value of N_u, say at $N_u = C_0$; this is an absorbing boundary, once reached it cannot be exited. Since the population never reaches the boundary $N_u = \infty$, it is strictly a natural boundary. However, we can treat it as a reflecting boundary. The boundary conditions are $T_{met}(C_0) = 0$ and $\partial_{N_0} T_{met}(\infty) = 0$. Solving Eq. 4.31, the MET is

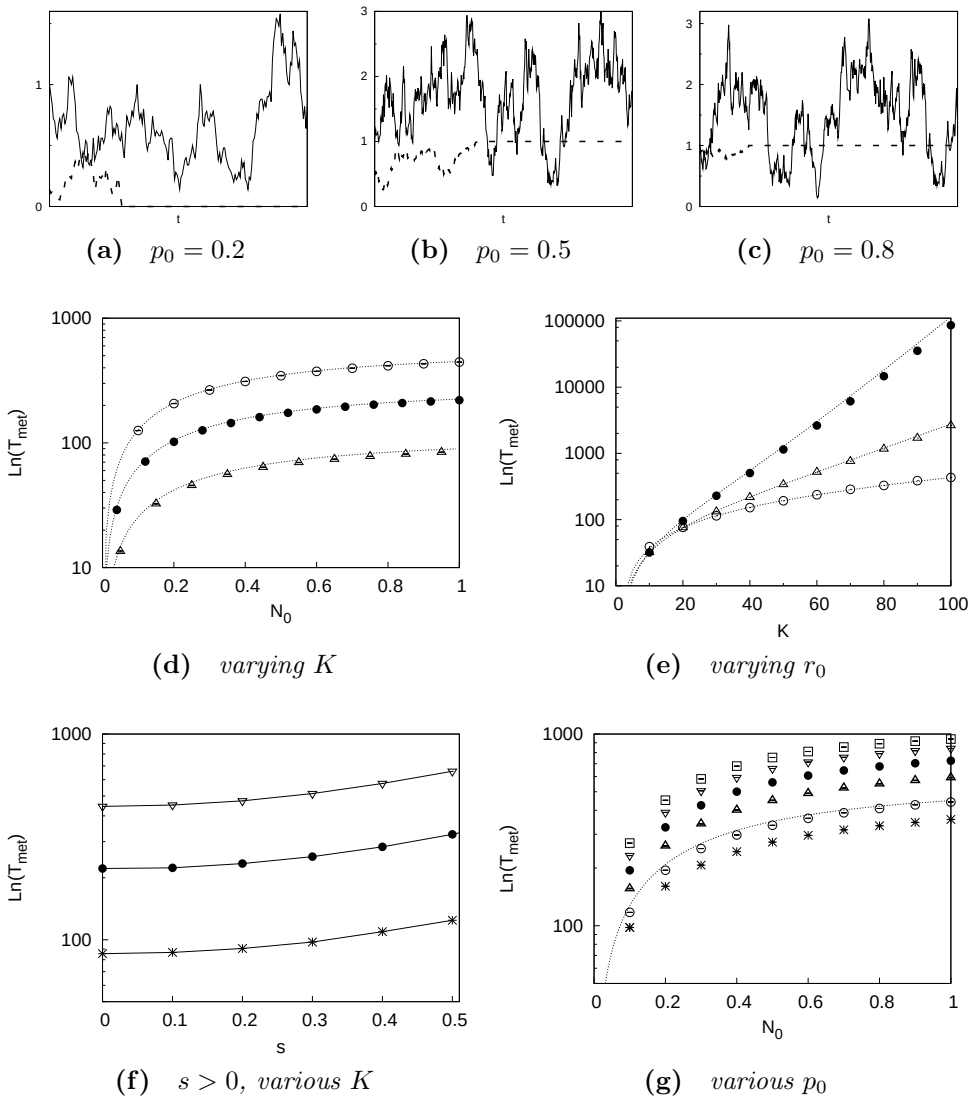
$$T_{met}(N_0) = \int_{C_0}^{\infty} G(N_u,\, N_0)\, \mathrm{d}N_u, \tag{4.32}$$

where $G(N_0, N_u)$ is the Green's function of the BFPE Karlin & Taylor (1975*b*),

$$G(N_u, N_0) = \begin{cases} \frac{2\Big(S(N_u) - S(C_0)\Big)}{s(N_u)B(N_u)} & \text{for } N_u < N_0 \\ \frac{2\Big(S(N_0) - S(C_0)\Big)}{s(N_u)B(N_u)} & \text{for } N_u > N_0 \end{cases} \tag{4.33}$$

where $s(x) = e^{-2\int^x \frac{A(y)}{B(y)}\,\mathrm{d}y}$ and $S(x) = -\frac{1}{2\mathbf{i}}\sqrt{\frac{\pi K}{r_0}} e^{-\zeta} \mathrm{Erf}(\mathbf{i}\sqrt{\frac{r_0}{K}}(K - x))$ (Erf $(\mathbf{i}z)$ is the error function with an imaginary argument, Abramowitz & Stegun 1972).

Simulation results for the mean extinction time measured in discrete generations, are shown in Fig. 4.10. (Note the figures actually show $\ln(T_{\mathrm{met}})$.) Above a critical size N_c, the population typically spends a long time fluctuating about the discrete quasi-stationary population size, N_*, (Lande et al. 2003, Lande 1993). Within this time, the population may become monomorphic, see figures 4.10a to 4.10c. Population trajectories that are initially close to the extinction boundary



(a) $p_0 = 0.2$ **(b)** $p_0 = 0.5$ **(c)** $p_0 = 0.8$

(d) *varying* K **(e)** *varying* r_0

(f) $s > 0$, *various* K **(g)** *various* p_0

Figure 4.10 *Mean extinction times (in the unscaled variables) for: top row, sample paths of* N *(solid line) and* p *(dashed line) for* $N_0 = K$, $s = 0.5$ *with (a)* $p_0 = 0.2$, *(b)* $p_0 = 0.5$ *and (c)* $p_0 = 0.8$. *Middle row, neural MET results for (d)* $K = 20, 50, 100$, *for varying* N_0, *(e)* $r_0 = 0.01, 0.05, 0.1$ *(bottom curve to top curve) for varying* K. *Continuous model results (dashed lines) and discrete simulation results (points) are shown. Bottom row, nonneutral MET results, (g)* $K = 20$, $s = 0.5$ *and initial frequencies* $p_0 = 0.1, 0.2, 0.4, 0.6, 0.8, 1.0$ *(bottom curve to top curve) for varying* N_0. *(f)* $N_0 = K$, $K = 20, 50, 100$ *(bottom curve to top curve) for varying* s. *Unless otherwise stated* $r_0 = \frac{1}{K}$. *All error bars are too small to be visible.*

are more likely to fall to zero before reaching N_*. This explains the rapid decline in the MET for increasingly small initial sizes in all the MET curves shown.

In the absence of selection, if K is large, the quasi-stationary population size is approximately the stable deterministic state, $N_* \sim K$. Crudely speaking, we expect that if the quasi-stationary population increases, it will take longer for extinction to occur. Clearly this holds the larger the carrying capacity, see Fig. 4.10d. However, the intrinsic growth rate will also influence the MET. At small carrying capacities, the MET appears to be unaffected by r_0, see Fig. 4.10e. However, for larger K, a larger intrinsic growth rate significantly delays extinction.

The analytical expression for the neutral MET in Eq. 4.32 is calculated numerically and is in excellent agreement with the simulation data. Notice however that for the largest value of r_0 in Fig. 4.10e, there is a slight discrepancy between theory and the data. This is because large r_0 violate the diffusion constraint $r_0 \sim O(\frac{1}{K})$, as explained in §4.4.4.

Once selection is included, the extinction time is sensitive to the strength of selection and the initial genetic variation in the population. In Fig. 4.10f the MET for different selective strengths is shown. The population is initially maximally polymorphic and at the intrinsic carrying capacity, $p_o = 0.5$, $N_0 = K$. Selection increases the extinction time, as shown in Fig. 4.10f. However, this critically depends on the initial state of the population. As shown in Fig. 4.10g, if initially there is not enough genetic variation in the population, that is p_0 is below a critical allele frequency p_c, the population is doomed to a speedy decline. This is clear from the two curves, for $p_0 = 0.1$ and 0.2, that fall below the neutral MET prediction in Fig. 4.10g.

Fixation probability

We now consider the fixation probability of the favoured allele using simulation data and results from previous studies.

In a population with initial size N_0, let $\pi_s(p_0, N_0)$ be the fixation probability of the favoured allele with selective advantage s and initial frequency p_0. For a constant effective population size N_e, the classic result for a neutral allele is (Crow & Kimura 1970)

$$\pi_0(p_0) = p_0. \tag{4.34}$$

If the selective advantage s_u is small, the fixation probability is (Kimura 1962,

Kimura & Ohta 1968),

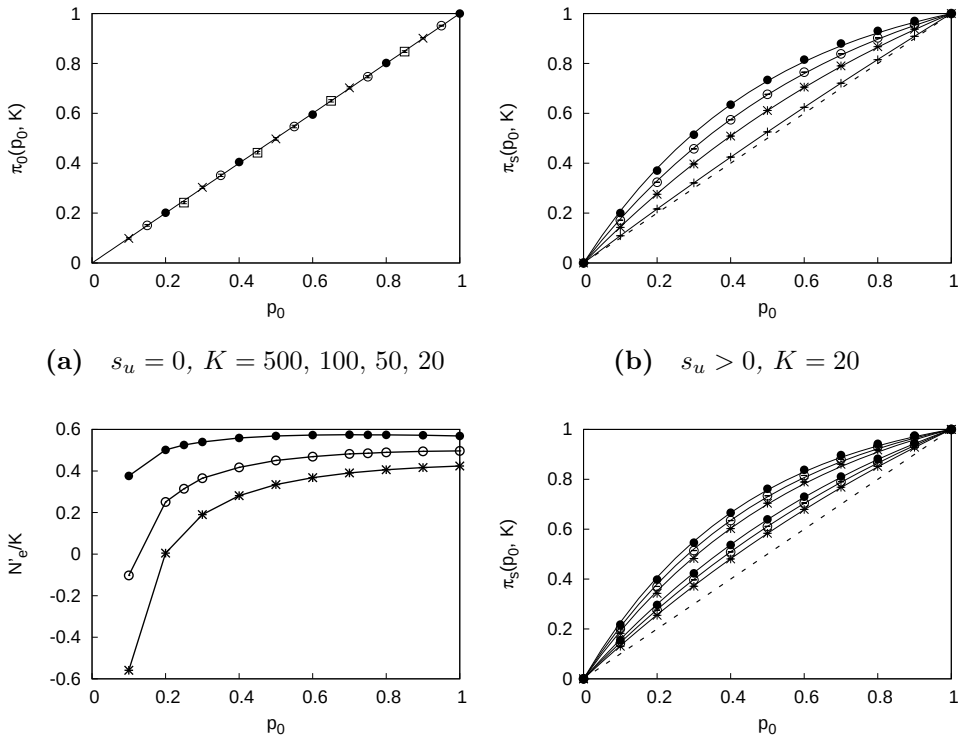
$$\pi_s(p_0,\, N_e) = \frac{1 - e^{-4N_e s_u p_0}}{1 - e^{-4N_e s_u}}. \tag{4.35}$$

If only one '+' allele is initially present in the local population and selection is weak, then the classic result for the fixation probability is $\pi_s \approx 2s_u$ (Haldane 1932, Fisher 1958), which was found using a branching process (Cox & Miller 1980).

Many studies have investigated fixation when the population size can change. For example, the population size is treated as cyclically varying or subject to a single, discontinuous change (Ewens 1967, Kimura & Ohta 1974, Otto & Whitlock 1997, Pollak 2000). In these studies, an effective population size that encapsulates the effect of these demographic variations on the fixation probability is determined. The effect of extinction and recolonization is the focus of metapopulation studies of fixation (Slatkin 1981, Lande 1985*b*, Barton 1993, Whitlock 2003, Vuilleumier et al. 2008, Aguileé et al. 2009). Recently, fixation in a population under environmental and demographic stochasticity (Lambert 2006, Parsons & Quince 2007*a*,*b*, Parsons et al. 2008, Engen et al. 2009) has been investigated (see also Karlin 1968, Heyde & Seneta 1975, Donelly & Weber 1985). The general observation is that the fixation probability of an advantageous allele increases with an expanding population, and decreases with a shrinking population size. This is confirmed analytically by Parsons et. al, who examine a two-dimensional logistic birth-death process in which two types of individuals with differential birth and death rates (hence different fitnesses) compete to survive (Parsons & Quince 2007*a*,*b*, Parsons et al. 2008).

Simulation data for the fixation probability in the MWFM are shown in Fig. 4.11 for different initial allele frequencies. The initial population size of the intrinsic carrying capacity is K (and $\zeta = 1$, so $r_0 = \frac{1}{K}$). If the local allele is selectively neutral, the fixation probability is identical to the constant population size result, $\pi_0(p_0,\, N_0) = p_0$, see Fig. 4.11a. This is expected, since in the neutral case, the stationary allele frequency distribution is identical to Wright's allele frequency distribution for constant population size, see Eq. 2.4. This makes sense because in the neutral case, the effect of demographic stochasticity is to simply enhance genetic drift. Hence, it is easy to see that the chance for a single mutant to fix is still $1/N_0$, which is the initial frequency of the mutant. Parsons et al. (2008) have demonstrated this result analytically for a density dependent birth-death process.

For the nonneutral case, simulation data are compared to the constant



(a) $s_u = 0$, $K = 500, 100, 50, 20$ **(b)** $s_u > 0$, $K = 20$

(c) $s_u > 0$, N_e **(d)** $K = 10, 20, 100$

Figure 4.11 *Fixation probability $\pi(p_0, N_0)$ with initial size $N_0 = K$, ($r_0 = \frac{1}{K}$). Simulation data (various symbols) have been fit to: (a) the neutral constant population size fixation probability, $\pi_0(p_0)$, for $K = 500, 100, 50, 20$ against. All symbols lie on the line. (b) the nonneutral constant population size fixation probability,$\pi_s(p_0)$, for $K = 20$ and $s_u/r_0 = 0.2, 0.5, 0.75, 1$ (from bottom curve above the straight line, to top curve). (c) the estimated effective population sizes N'_e scaled by K for $K = 20, 50, 100$ (from the bottom curve to the top). (d) fixation probability for $s = 0.5$ (bottom curve set) and $s_u/r_0 = 1$ (top curve set) for $K = 10, 20, 100$ (bottom to top) in each set of curves.*

population size result, Eq. 4.34, see Fig. 4.11b and an effective population is determined for each fixation probability curve. The effective population sizes that give the best fit are shown in Fig. 4.11c. As we saw in the last section, a typical stochastic trajectory the population spends an exponentially long time (in the system size ζ) fluctuating about the quasi-stationary population size, N_*. To zeroth order in ζ^{-1}, this will be close to the deterministic stable population size, $N_d = K^2\overline{r}$. Typically, the population becomes monomorphic before extinction occurs (Parsons & Quince 2007*a*). Therefore, the long-time, average effect of demographic stochasticity is to augment genetic drift and this is captured by the effective population size N_e. Since the quasi-stationary population size increases with increasing selection, the fixation probability increases with selection. Figure 4.11c shows the fixation probability curves for different K for two different values of s. Notice that for each s, the larger the carrying capacity, the greater the fixation probability.

The fit between simulations and $\pi_s(p_0)$ is surprisingly good even for carrying capacities as low as $K = 10$, see Fig. 4.11c; though the intrinsic carrying capacity is low, the quasi-stationary population size may be large enough to suppress fluctuations in N.

The estimated effective population sizes, N_e', that give the best fit to the data for different carrying capacities are shown in Fig. 4.11d. The effective population size is on the order of K and varies with s. The estimated sizes appear to converge to a fixed value as the initial frequency approaches 1. Negative values of N_e' indicate that the local allele behaves as if it is deleterious. This occurs for low initial frequencies, when the '-' allele is common.

Transition rates

In this section we determine the rate at which the population transitions between the two ecological zones. Transitions between alternate states can drive adaptation (Haldane 1932, Wright 1940, 1968-78). Wright was one of the first to recognise this: his "shifting balance" theory describes transitions between adaptive states as a means for complex adaptation (Wright 1968-78). Wright believed that genetic drift could cause a population to stochastically shift between adaptive peaks (Wright 1931, 1932). Stochastic peak shifts have been widely studied in the literature. For example, under different forms of selection and trait structures (Lande 1986, Barton & Rouhani 1987*a*,*a*, Lande 1985*c*), large populations (Goodnight 2005), spatially distributed populations (Barton & Rouhani 1987*b*, Burton & Travis 2008; see also Barton & Charlesworth 1984,

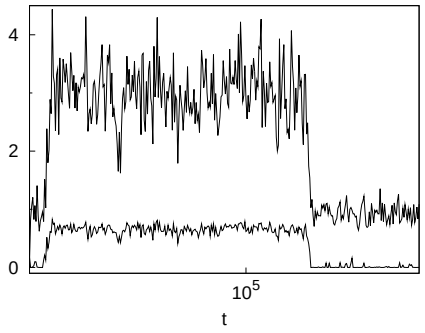
Newman et al. 1985) and deterministically expanding populations (Rouhani & Barton 1987). Peak shifts can also be caused by mechanisms other than genetic drift, such as changes to the adaptive landscape (Fisher 1958, Whitlock 1997, Hadany 2002).

Peak or zone shifts due to demographic and genetic drift have not previously been considered. Whilst both zones "exist", the population can shift out of its ecological niche into a state of maladaptation. The frequency with which the population does this (and the reverse) sets the rate of local adaptation. Here we determine an upper bound for the rate of transitions between zones, comparing these bounds with simulation data.

In the limit of weak noise, the stochastic dynamics can be separated into two timescales. On the short timescale, the population is trapped within the domain of attraction of one of the two zones, fluctuating about its corresponding peak. What can happen on longer timescales is escape from that zone: after an exponentially long time, stochastic perturbations will eventually carry the population across the unstable saddle connecting the two zones, to the alternative ecological zone. Such "jumps" between states are rare. However, though the time between shifts is exponentially long, the shifts themselves happen rapidly (Van Kampen 2007, Lande 1985*a*), as can be seen in Fig. 4.12a; the time series show a striking "punctuational" pattern, which has been previously observed by several authors (Newman et al. 1985, Lande 1985*a*, Holt et al. 2003) .

The transition rate can be defined and calculated in several different ways (Caroli et al. 1980, Hänggi 1986, Hänggi et al. 1990, Risken 1996). It can be defined as the number of trajectories that reach the neighbouring domain for the first time, normalized by the total number of trajectories leaving the domain of origin (see Kramer's rate and flux-over-population in (Hänggi et al. 1990)). Alternatively, it could also be defined as an average transition rate, which is the inverse of the average time to shift from one zone to the other. The latter definition invites use of the mean first passage time (MFPT) as the inverse transition rate (Gardiner 1985, Van Kampen 2007, Talkner 1987, Hänggi et al. 1990). Since we only seek an approximate upper bound to the transition rate, and because of its elegant simplicity, the latter is our method of choice (see also Risken 1996, Ch. 5). In the limit of weak noise, and after long times the two approaches should give equivalent results (Boilley et al. 2004, Reimann et al. 1999).

Let Γ_{jk} denote the transition rate from zone j to zone k. We need to know the possible paths the population may take through state space. This depends



(a)

Figure 4.12 *Sample paths of stochastic trajectories for population size N (upper curve) and allele frequency p (lower curve). The paths are sampled every 100 generations, with $s = 8$, $\hat{p} = 0.002$, $M = 16$, $K = 500$ and $\zeta = 1$. The (scaled) population size has been halved for comparison with the allele frequency path. Notice that N and p shift at approximately the same time.*

on the saddle geometry and the dynamics. We know that there is only one (simple) saddle near the separatrix. Moreover, we know from detailed balance that the most likely path from zone j to domain k is the precise opposite of the reverse transition (Barton & Rouhani 1987*a*). This eliminates the possibility of more complicated paths arising from limit cycles or chaotic attractors. A typical population trajectory may be as follows. Initially the population fluctuates within domain j, spending a long time there. Eventually the population will reach the boundary of domain j, the separatrix, and cross into the saddle (we assume the boundaries of the saddle and separatrix approximately coincide. This will be accurate for low drift: compare Eq. 4.15 in the limit $\zeta \to 0$ with Eq. 3.12). The average time it takes for this to happen is the mean first passage time, τ_j. Once at the saddle the population will either (*i*) return to domain j or (*ii*) dither for some time at the saddle. If the former happens then this trajectory will not count towards the Γ_{jk}. If, however, the population moves on from the saddle to domain k, this trajectory will count towards the transition rate. The longer the trajectory takes to reach domain k, the smaller the transition rate. Hence an upper bound on Γ_{jk} would be the transition rate from domain i to the saddle ('0'), Γ_{i0}. The upper bound will be close to the actual transition rate if there are few return trajectories from the saddle bask to domain j, and also if the population spends little time at the saddle. For systems with potential dynamics and weak noise, a simple saddle geometry usually implies that the time spent traversing the saddle is orders of magnitude smaller than the MFPT and hence can be neglected (e.g. Lande 1985*a*; see also Boilley et al. 2004).

With these conditions in place, the upper bound on the rate of transitions

from the ecological niche to the maladaptive zone, Γ_{+-}, is equivalent to the rate of escape from ecological zone to the saddle, Γ_{+0}, conditional on absorption at the separatrix,

$$\Gamma_{+-} \leq \Gamma_{+0}. \tag{4.36}$$

To determine the upper bound we need only find the MFPT. Generally, multidimensional mean exit times are too difficult to solve exactly and asymptotic methods must be used. Under weak noise we can use the result derived by Talkner (1987) for τ_+, which is also reviewed in Hänggi et al. (1990). The derivation of this result is reproduced in App. B. Strictly, the MFPT is dependent on the initial location of the population within the zone of origin. For weak noise however, provided the initial location is far from the immediate neighbourhood of the saddle, the MFPT is constant across the entire domain (Gardiner 1985). Then rate Γ_{+0} is simply the inverse of τ_+. Once at the saddle, the probability of crossing the saddle is taken to be about one half, this gives (Hänggi et al. 1990, Eq. 2.33),

$$\Gamma_{+0} = \frac{1}{2\tau_+}. \tag{4.37}$$

Taking the probability at the saddle to be $\frac{1}{2}$ also greatly overestimates the flux across the saddle, since the population may stall at the saddle, crossing and recrossing multiple times (Barton & Rouhani 1987*a*). Talkner found that for a general multidimensional diffusion process with a sharply-peaked stationary distribution, the escape rate from a domain j to the absorbing separatrix is (Talkner 1987, Eq. 4.26),

$$\Gamma_{j0} = \frac{\nu_0}{2\pi}\Big(\frac{P(0)}{P(j)}\Big), \tag{4.38}$$

where $P(k)$ is the equilibrium peak probability for zone k. Variable ν_0 is the leading eigenvalue of the Jacobian of the deterministic dynamics evaluated at the saddle. Evaluating the Jacobian at the saddle (to find ν_0) is approximately the same as evaluating the Jacobian at the unstable deterministic state in the limit of low drift (i.e. at the separatrix, see App. B for details). Using the zero-width peak approximation and noting that ν_0 can be found from matrix Ω (which is related to the Jacobian, see App. B, the section titled 'Matrices $\mathbf{\Omega}$ and $\mathbf{\Sigma}$'), we obtain

$$\Gamma_{+-} \leq \frac{\nu_0}{2\pi}\Big(\frac{N_+V_+}{N_0V_0}\Big)e^{-2(U_+-U_0)}, \tag{4.39}$$

$$\Gamma_{-+} \leq \frac{\nu_0}{2\pi}\Big(\frac{N_-V_-}{N_0V_0}\Big)e^{-2(U_--U_0)}, \tag{4.40}$$

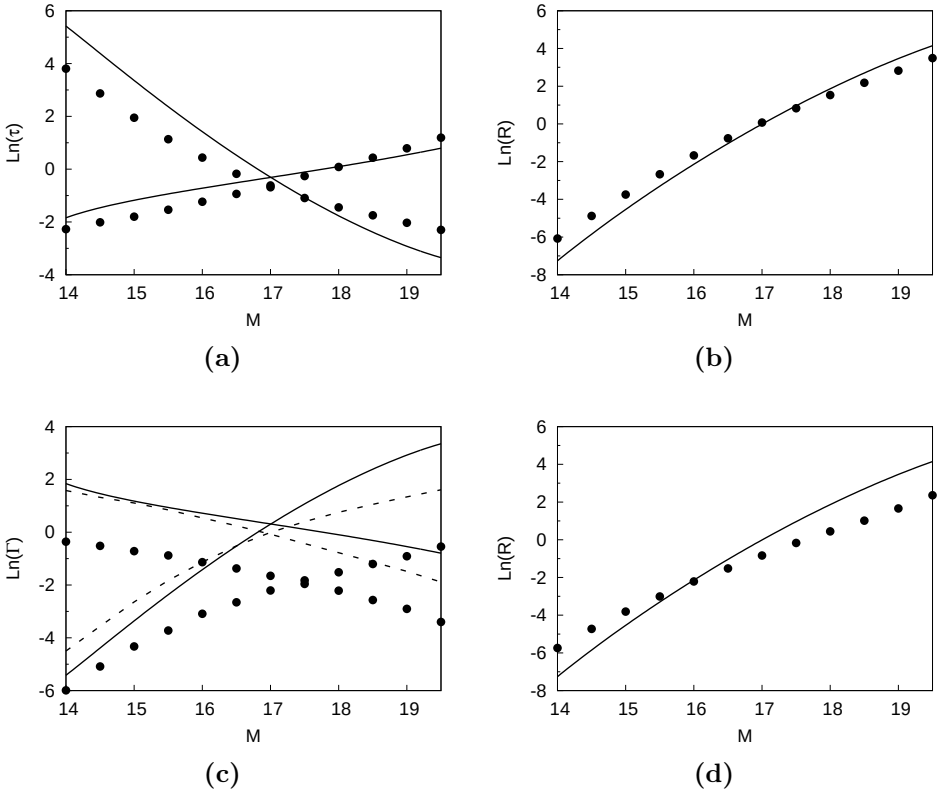


Figure 4.13 *The logarithm of the mean exit times* $\ln(\tau_j)$ *and transition rates* $\ln(\Gamma_{j0})$ *for* $s = 8$, $\hat{p} = 0.002$, $\zeta = 1$ *(and* $K = 500$ *in simulations). Solid lines are numerical results. Two types of simulations were carried out: dots represent direct simulation of the transition rate, where the final transition rate was calculated as the average transition rate per run. The dashed lines represent simulation results of the inverse mean exit time. (a)* $\ln(\tau_+)$ *(increasing curves),* $\ln(\tau_-)$ *(decreasing curves) and (b)* $\ln R = \ln\Big(\frac{(\tau_+)^{-1}}{(\tau_-)^{-1}}\Big) \equiv \ln\Big(\frac{\Gamma_{+0}}{\Gamma_{-0}}\Big)$. *(c)* $\ln(\Gamma_{+0})$ *(increasing curves),* $\ln(\Gamma_{-0})$ *(decreasing curves) and (d)* $R = \ln\Big(\frac{\Gamma_{+0}}{\Gamma_{-0}}\Big)$.

We expect this to be accurate when the potential difference $|U_0 - U_k|$ is large and drift is weak. The dominating factor in the rate calculation is the exponential factor involving the potential barrier; this is the barrier the population must overcome in order to reach the alternate zone. This compares with the exponential dependence of the mean time to shift on the product of the population size in constant population size models (Barton & Charlesworth 1984, Lande 1985*a*).

The rates are proportional to the leading eigenvalue at the saddle ν_0, which characterises the degree of instability at the saddle. The standard form for rate expressions usually involve an 'entropy' factor (Weiss 1986, Hänggi 1986, Barton & Rouhani 1987*a*), that is, the ratio of the determinant at the saddle to the determinant at the stable domain. It is in fact the ratio between the number of states at the ecological zone to the number of states at the saddle: a wider saddle would increase the frequency of shifts because it would contain more states. The term arises from assuming the stationary peaks are Gaussian. Had we used the Gaussian approximation, the term $\sqrt{\left(\frac{\det(\mathbf{\Omega})_{\mathbf{x}_j}}{|\det(\mathbf{\Omega})|_{\mathbf{x}_0}|}\right)}$ would multiply the rate expressions.

In Fig. 4.13 the theoretical mean exit times and transition rates are compared to simulation data. Note that the logarithm of the rates and mean exit times are shown. We compare two different discrete rates, the average "per run rate" and the inverse of the mean exit time. The former was obtained as follows. The population was set to be initially far from the saddle. The per run rate was defined as the inverse of the number of generations it took the population to reach the saddle. The final transition rates shown here are averages of the per run rate. The mean exit time was found by counting the number of times a domain was visited. The time spent within each domain was added to a total 'domain time' count. The final mean exit times were calculated as the total domain time divided by the number of domain visits. Simulation runs were on the order of 10^5 runs. The saddle was set to be the deterministic unstable state. Treating the saddle as a single threshold instead of a range of states introduces a slight error into the simulation data, which is a negligible contribution to the large discrepancy between theory and simulations. Moreover, in using the zero-width approximation, the shape of the saddle does not feature in the expression for the rates.

The discrete mean exit times qualitatively agree with the theoretical expression for τ_j, from Eq. 4.37: as migration increases, the average time to leave the maladaptive zone increases and vice versa for the ecological niche, see Fig. 4.13a. This is because the probability of finding the population in the maladaptive zone increases with M. The qualitative agreement may be attributed to the that fact the peak probabilities feature in the theoretical expression and is evident from taking the ratio of the inverse mean exit times, see Fig. 4.13b. As was shown in §4.3.2, the peak probabilities from simulations agree qualitatively with theory, (once we take into account the inherent error between the theoretical and discrete dynamics, see §4.4.4). However, the theory fails to predict the correct

mean exit times, which are least one order of magnitude out. Since theoretical mean exit times are not upper bounds, the quantitative agreement is quite poor. The main source of error for this discrepancy is most likely due to the saddle and in particular the exponential potential term. The peaks are not well separated for $M > 18$ and though the saddle is visible, the potential depth between the high state peak and the saddle is small: the asymptotic result is expected to do poorly when the potential depth is small.

Similarly, the discrete transition rates qualitatively agree with theory: as migration increases, the transitions out of the maladaptive zone become less frequent, see Fig. 4.13c. The converse is true for the ecological niche. Comparing theory with the average per rate run, the latter is several orders of magnitude smaller. We expect the quantitative agreement to be better than this since we simulated the transition rate up until the saddle and not the transition between domains. Since the discrepancy between theory and simulation is smaller for the logarithm of ratio $R = \frac{\Gamma_{+0}}{\Gamma_{-0}}$ (see Fig. 4.13d), it seems either the assumptions that the calculation are based upon are invalid for our model, and/or the continuous model is not the correct diffusion limit of the discrete model. We suspect that some error will arise from the latter, as we have already seen when comparing the zone probabilities and stationary moments in §4.3.2 and §4.3.2. However, this error cannot account for the orders of magnitude discrepancy since the stationary properties of the discrete model are reasonably well approximated by the continuous theory, e.g. ratio R. Therefore, perhaps the error is due to the assumptions we have based the transition rate calculation on. In any case, the transition rates need further investigation (we continue this discussion in §4.4.4, where we discuss the overall success of the continuous theory with respect to the discrete simulation results).

This concludes our investigation of the single locus model. We have shown that for a given range of parameters, when gene flow is highly maladaptive, the population fluctuates for long times in a state of maladaptation or between adaptive states within its ecological niche. We have investigated the mean time until extinction, fixation probability and shift rates. Further work on these first-passageproperties would be worthwhile.

We now consider the multilocus model.

4.4 Multiple loci

What difference do multiple loci make to local adaptation and persistence once random drift is taken into account? Can the population still be found in one of two ecological zones? To answer these question we concentrate on the stationary properties of the multilocus model, carrying out a similar analysis to the single locus case. We then briefly consider the rate of transitions between zones.

4.4.1 Stochastic equilibrium

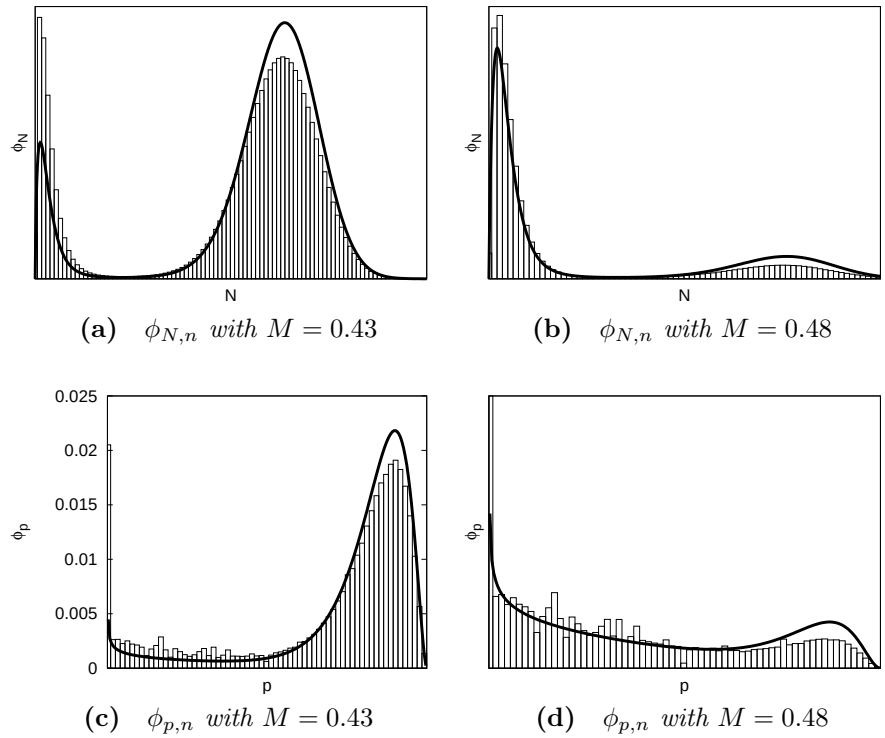
The behaviour of the population at stochastic equilibrium is qualitatively similar to the single locus case. In the last chapter, it was found that at equilibrium, bistable loci simultaneously occupy their respective high states or their respective low states. Therefore, there can be at most two stable attractors. Moreover, if loci are symmetric and bistable, there is only one unstable (biologically valid) state, since loci have the same equilibrium frequency (see §3.4).

In the stochastic model, when selection and migration dominate drift, the multilocus stationary distribution, $\Psi_n(\mathbf{x})$, also becomes bimodal; the population can find itself fluctuating within a well-adapted or maladapted ecological zone. The multilocus distribution Ψ_n has already been derived in §4.2.1 and is given in Eq. 4.7. The distribution may be multi-peaked for certain parameter values that cause 'ghost peaks' to appear at the boundaries. However, we are only interested in the bimodality of Ψ_n, the conditions for which the two ecological zones exist and how these contrast with the single locus case. Unless otherwise stated, all loci are symmetric, i.e. allelic effects $\alpha_j = 1$ and migrant frequency $\hat{p}_j = \hat{p}$ for all loci. This is a restrictive assumption, but one that simplifies calculations. The effect of small asymmetries in α and $\hat{p}$ will be hard to distinguish from the effect of fluctuations.

Marginal distributions

We use the marginal distributions of population size, $\phi_{N,n}$, and allele frequency, $\phi_{p,n}$, to illustrate bimodality in Ψ_n. These distributions can be determined by integrating out the unwanted variables from Ψ_n, as was done for the single locus case in §4.3.1. For $\phi_{N,n}$, we integrate Ψ_n over all allele frequencies, similar to the integral for ϕ_N (see Eq. 4.9), which gives

$$\phi_{N,n} = \frac{1}{Z} N^{2M\zeta - 1} e^{2N\zeta \overline{r}_{min} - \zeta N^2} Y_n. \tag{4.41}$$



(a) $\phi_{N,n}$ *with* $M = 0.43$ **(b)** $\phi_{N,n}$ *with* $M = 0.48$

(c) $\phi_{p,n}$ *with* $M = 0.43$ **(d)** $\phi_{p,n}$ *with* $M = 0.48$

Figure 4.14 *Marginal distributions* $\phi_{N,n}$ *(top) and* $\phi_{p,n}$ *(bottom) for 10 symmetric loci with* $s = 0.5$, $\hat{p} = 0.2$ *and* $\zeta = 2$ *for* $M = 0.43$ *(left) and* $M = 0.48$ *(right). Solid curves are theoretical results, bars are simulation data.*

where $\overline{r}_{min} = 1 - ns\eta$ and $Y_n = \left({}_1F_1(4M\zeta\eta\hat{p}, 4M\zeta\eta; 4N\zeta\eta s)\right)^n$. Similarly for $\phi_{p,n}$, we integrate Ψ_n over N (see Eq. 4.12) to get,

$$\phi_{p,n} = \frac{1}{Z} p^{4M\zeta\hat{p}-1} q^{4M\zeta\hat{q}-1} I(p), \tag{4.42}$$

where

$$I(p) = \int_0^\infty e^{4N\zeta\eta sp} N^{2M\zeta-1} e^{2N\zeta\overline{r}_{min}-\zeta N^2} Y_{n-1}\, \mathrm{d}N. \tag{4.43}$$

The normalisation constants and integral I can be calculated numerically.

The main differences between the single locus and multilocus case in the deterministic model carry over to the stochastic model. These are highlighted by the marginal distributions shown in Fig. 4.14 for two different migration rates. The two ecological zones can now make an appearance for weak selection, low migration and polymorphic migrants when more than one locus contributes to fitness. This is because though the selection coefficient per locus is small, the 'effective' selection coefficient, $ns\eta$, which contributes to the average growth rate, can be large, (see §3.4). Notice the boundary peak of the maladapted zone in

$\phi_{p,n}$ and its corresponding 'interior' peak in $\phi_{N,n}$.

Bimodality in haploid stationary distributions ($\eta = \frac{1}{2}$, not shown) is not simply a consequence of 'ghost peaks' as it is for the single locus case, but is now due to the underlying deterministic dynamics. Because of fluctuations, gene flow also acquires an 'effective' rate $M\eta$; ploidy now makes a nontrivial difference to the stochastic population and we cannot simply treat haploid results as 'scaled' diploid results, as is possible in the deterministic model.

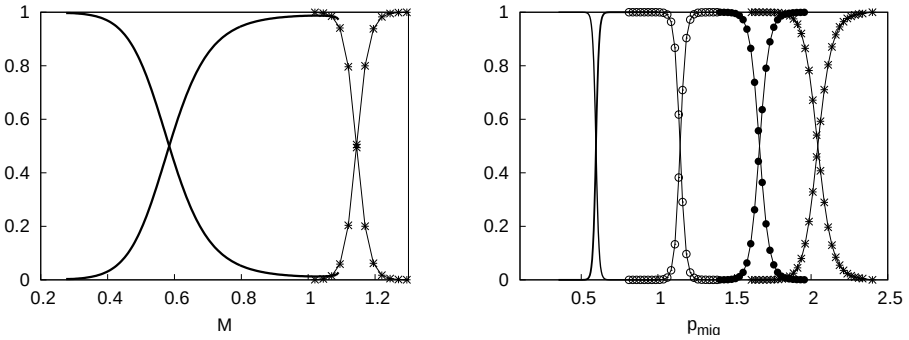
We next determine the probability of finding the population within a given zone, followed by the stationary moments of N and p, under the assumption of low drift.

4.4.2 Zone probabilities

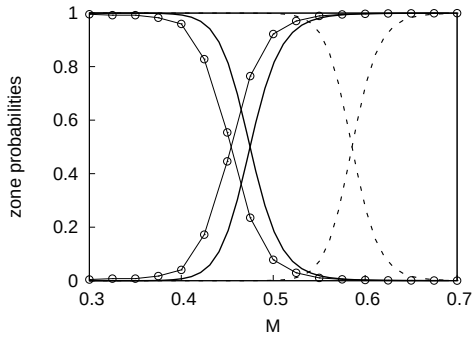
The probability $P(k)$ of finding the population within zone k is the integral over Ψ_n about the k^{th} peak. Using the local peak approximation (§4.3.2) to approximate the multilocus stationary distribution, the approximate peak probabilities are given by Eq. 4.25, once the single locus distributions and weights are replaced with the appropriate multilocus distributions and weights. We use the Gaussian approximation throughout, $\tilde{\Psi}_{G,n}$, unless otherwise stated (see App. B for details). Note that the population is within an ecological zone when the allele frequency of *all* loci fluctuate within the domain.

In Fig. 4.15, probabilities $P(+)$ and $P(-)$ are shown. As the number of loci increases, higher migration is required for maladaptation to be probable. This is clear from Fig. 4.15a; rate M_{max} for which $P(+) = P(-) = \frac{1}{2}$, increases. Notice that the range of migration rates over which the peak probabilities transition decreases and the transition becomes sharp; that is, the bimodal range decreases the greater the number of loci. The more common the local allele is on the mainland, $\hat{p} \gg 0$, the larger migration has to be for maladaptation to be probable, see Fig. 4.15b. Additionally, the bimodal range increases; even though maladaptation requires high migration rates, once maladaptation becomes possible, the population remains vulnerable over a greater range of rates. These features mirror the underlying deterministic behaviour of the population.

In Fig. 4.15c, the peak probabilities for ten loci are shown, with moderately weak selection, ($s = 0.5$, $\zeta s = 1$), and migration is low, ($\zeta M \sim 1$). The agreement between simulation data from the MFWM and the Gaussian approximation is reasonable, though the latter once again estimates a slightly higher M_{max}; this is not due to a numerical error but is due to the discrete dynamics (see §4.4.4). The Gaussian approximation scheme does better than the zero-width approximation



(a) *varying* n **(b)** *varying* $\hat{p}$



(c) $s = 0.5$, $\hat{p} = 0.2$, $\zeta = 2$

Figure 4.15 *Multilocus zone probabilities* $P(+)$, $P(-)$ *for (a)* 4 *loci (solid lines), 10 loci (with crosses), (b)* 10 *loci with migrant frequencies from left to right,* $\hat{p} = 0.001, 0.1, 0.2$ *and* 0.25 *and (c) ten loci with* $s = 0.5$, $\hat{p} = 0.2$ *and* $\zeta = 2$. *Approximations using* $\tilde{\Psi}_{G,n}$ *(solid lines), simulation data (with open circles) from the MWFM for* $K = 200$, $r_0 = 0.01$ *and approximations using* $\tilde{\Psi}_{0,n}$ *(dashed lines) are shown. Unless otherwise stated,* $s = 1$, $\hat{p} = 10^{-4}$ *and* $\zeta = 1$ *for diploids.*

of $\tilde{\Psi}_n$.

4.4.3 Stationary moments

The stationary moments approximated by Eq. 4.26, derived in §4.3.2, are applicable to the multilocus case by once again replacing the appropriate single locus quantities with their multilocus counterparts.

Figure 4.16 shows numerical calculations of the exact demographic and genetic stationary moments when ten loci contribute to fitness. We saw in the last section that the transition between niche-dominance and maladaptation requires higher migration rates the more polymorphic immigrants are or the stronger selection is. The same trends are seen in the averages of N and p, see figures 4.16a, 4.16c, 4.16e and 4.16g. The more common the local allele, the smaller the drop in the average population size and the average allele frequency. The inverse trend is

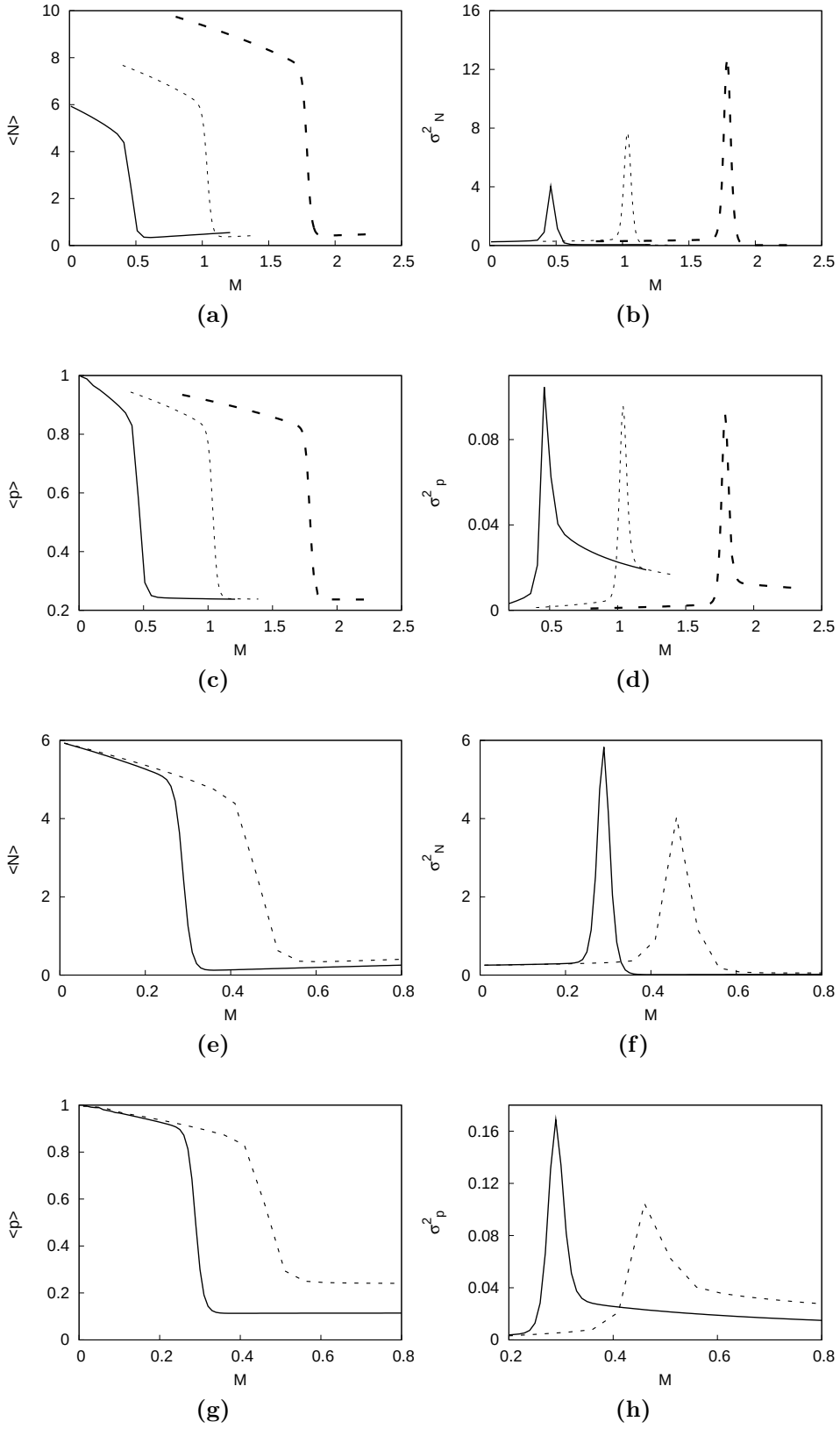


Figure 4.16 *Stationary moments for ten loci with varying s,$\hat{p}$ (e)-(h). In figures (a)-(d), s varies: from the leftmost curve moving rightwards, $s = 0.5, 0.75, 1$. In figures (e)-(h) $\hat{p}$ varies: from the leftmost curve $\hat{p} = 0.1, 0.2$. Unless stated otherwise, $s = 0.5$, $\hat{p} = 0.2$ and $\zeta = 2$.*

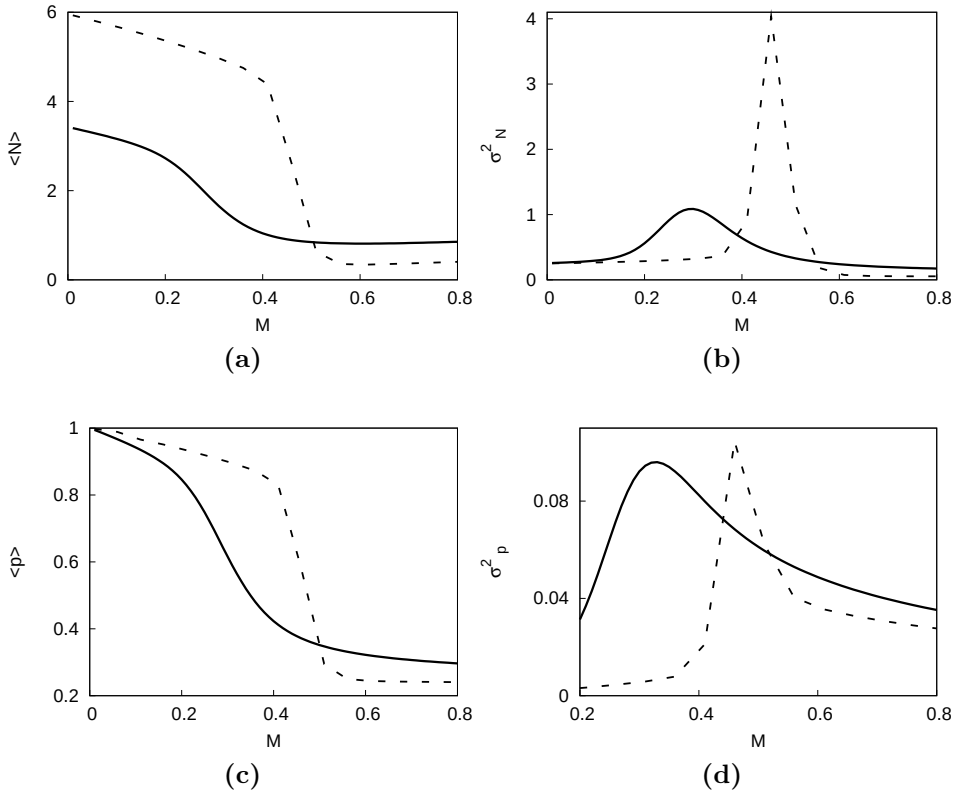


Figure 4.17 *Stationary moments of N (top) and p (bottom) for 5 loci (solid lines) and 10 loci (dashed lines) with $s = 0.5$, $\hat{p} = 0.2$ and $\zeta = 2$.*

true for stronger selection; the drop in average population size becomes larger for higher migration rates.

All variances peak at a critical rate M_{max}. For strong selection, demographic variation increases with increasing s, because the selective strength is tied up with growth (hard selection), see Fig. 4.16b. Additionally, strong selection eliminates genetic variation from the local population, as can be seen by the decreasing peak height of σ_p^2. The curiously large forward tail in σ_p^2 suggests a slow but steady decline in the genetic variation once the population is in the maladapted state, because selection is strong, see Fig. 4.16c. The less rare the local allele, the smaller the variation in population size and allele frequency, see figures 4.16f and 4.16h.

In Fig. 4.17, the stationary moments for different loci numbers are contrasted. The drop in values of the averages is steeper over the bimodal migration rates the larger the number of loci, see figures 4.17a and 4.17c. Moreover, the average population is proportional to the number of loci, because, all else being equal, the population's response to selection is greater (i.e. the effective selection coefficient is greater) the larger the number of loci, as in the deterministic model. Though

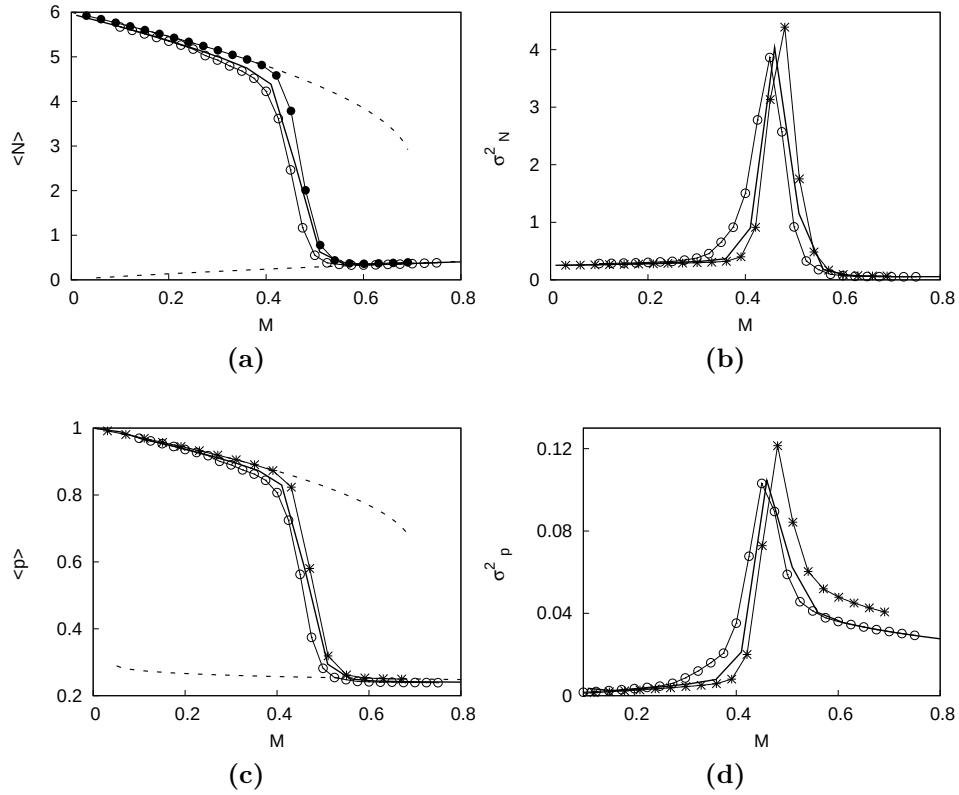


Figure 4.18 *Stationary moments of N (top) and p (bottom) for ten diploid loci with $s = 0.5$, $\hat{p} = 0.2$ and $\zeta = 2$. Numerical results (solid lines), simulation data (with open circles) for $K = 200\,, r_0 = 0.01$, and approximations using $\tilde{\Psi}_{G,n}$ (with crosses) are shown.*

the demographic and genetic variances are higher for larger n, variation is spread across a smaller range of migration rates than a smaller number of loci, see figures 4.17b and 4.17d.

Numerical calculations of the exact stationary moments are compared with simulation data and approximations in Fig. 4.18. The agreement between the three different types of results is reasonable. Notice that the averages are very close to the deterministic equilibria for migration rates above and below the bimodal range, see figures 4.18a and 4.18c. However the Gaussian approximation predicts a higher variance than is the case for either the numerical or discrete results.

We have examined the zone probabilities and stationary moments for the multilocus case. We next summarise what these results imply for the bimodal range, before moving on to the transient dynamics.

Significance of the bimodal range

There is no easy way to predict what the bimodal range, $M_{max} \pm \delta M$, will be for a given set of parameters. Nevertheless we can infer some simple trends from the results of the previous sections.

The stronger selection is per locus, the greater the number of loci and the higher the migrant polymorphism, the higher the critical migration rate M_{max} will be. Migration then needs to be high, $M \sim M_{max}$ for maladaptation to be more probable than adaptation. However, there is a trade-off, because strong selection and large loci numbers also reduce δM, and with it the population's chance to escape maladaptation; the transition between zones becomes sharp and their coexistence is fleeting. This may or may not be compensated for by higher migrant polymorphism, for which δM increases.

We have found that the stationary properties of the multilocus case are qualitatively similar to the single locus case, and that differences that arise are primarily due to the underlying deterministic dynamics. The multilocus mean exit times and transition rates can in principle be determined in the same way they were for the single locus case. However, we refrain from delving into these calculations here because of the discrepancies encountered in the single locus case.

Before moving on to examine the quantitative trait model, we discuss the discrete models, simulation results, and in particular how well the continuous theory approximates the discrete models.

4.4.4 Simulations

The multivariate diffusion process we have investigated is of interest in its own right, as a continuous model of a population undergoing gradual growth and evolution. We can also view the model as an approximation of an infinite number of similar discrete models in the limit of a large, or 'system size', parameter. Discrete models are more realistic than their continuous counterparts; the diffusion limit is also often taken as a means to circumvent mathematical difficulties (Ewens 1979). It is important, therefore, to check the validity of theoretical results against results derived from discrete models.

In this section we first describe the two discrete models we investigated for the single locus and multilocus cases: a Markov chain (MC) for population size under neutrality and the Modified Wright-Fisher model (MWFM) model. The

MFWM reduces to the MC under neutrality. We then discuss how well our theory approximates the discrete models. Note the models are described below in terms of the unscaled parameters (see §3.2.1). This should be obvious from the presence of r_0 and K in the expressions below, therefore for convenience we drop the subscript 'u'. The agreement between the discrete models and theory is discussed towards the end of this section.

The Markov chain (MC)

Under neutrality, for a single locus, the dynamics of population size decouple from genetic evolution and a Markov chain (Karlin & Taylor 1975*a*, Cox & Miller 1980) can be constructed for N: every generation a Poisson number of offspring are produced by equally fit parents. The mean of an individual's offspring distribution is $r_0(1 - \frac{N_t}{K})$, where N_t is the size of the t^{th} generation. After reproduction, a Poisson number of migrants with mean M join the local population. The transition probability of going from state $N_t = i$ to a state $N_{t+1} = j$ in the next generation is

$$p_{ij} = \frac{e^{-\theta_i}(\theta_i)^j}{j!}, \tag{4.44}$$

where $\theta_i = i + r_0 i(1 - \frac{i}{K}) + M$. Note that $r_0 i(1 - \frac{i}{K})$ is the logistic map, which can become chaotic if $r_0 \geq 2$ (May 1974, 1976, May & Oster 1976). We denote a Poisson random number from a Poisson distribution with mean θ as Poisson(θ).

Exact results for the stationary distribution or stationary moments are not obtainable. However, we know the stationary distribution $\mathbf{p}(\infty)$ satisfies $\mathbf{p}(\infty) = \mathbf{P}\mathbf{p}(\infty)$, which can be found numerically by iterating forwards in time from an initial state vector $\mathbf{p}(0)$ using the transition matrix $\mathbf{P}$

$$\mathbf{p}(t) = \mathbf{P}^t\mathbf{p}(0) \tag{4.45}$$

since $\lim_{t\to\infty} \mathbf{p}(t) = \mathbf{p}(\infty)$ (Karlin & Taylor 1975*b*, Ewens 1979).

The discrete chain can be approximated as a continuous diffusion process of transformed variables X (scaled population size) and T (scaled time), with infinitesimal mean and variance (Karlin & Taylor 1975*b*, Ewens 1979, Lande et al. 2003),

$$A_{mc} = \lim_{\triangle T \to 0} \frac{1}{\triangle T} \mathrm{E}[\delta \mathrm{X} | \mathrm{X}] \tag{4.46}$$

$$B_{mc} = \lim_{\triangle T \to 0} \frac{1}{\triangle T} \mathrm{V}[\delta \mathrm{X} | \mathrm{X}] \tag{4.47}$$

To find these moments, the variables are transformed to a scale where changes in N and t are small enough so that higher infinitesimal moments of X are negligible in the limit $\triangle T \to 0$,

$$\lim_{\triangle T \to 0} \frac{1}{\triangle T} \mathrm{E}[(\delta \mathrm{X})^{\mathrm{h}} | \mathrm{X}] = 0 \qquad\qquad h > 2. \tag{4.48}$$

The discrete variables are rescaled relative to the intrinsic carrying capacity K, the 'large' parameter in the system, $X = \frac{N}{K}$ and $\triangle T = \frac{1}{K}$. In the limit $K \to \infty$, the diffusion approximation becomes accurate (the limit $\triangle T \to 0$ is equivalent to $K \to \infty$). The infinitesimal moments are

$$A_{mc} = r_0 X(1 - X) + M, \qquad \text{and} \tag{4.49}$$

$$B_{mc} = X, \tag{4.50}$$

provided $r_0 \sim O(\frac{1}{K})$ and $M = O(1)$. This is equivalent to setting scaling parameter $\zeta = 1$. These moments describe the diffusion that best approximates the MC and is exactly equivalent to the continuous model under neutrality. Therefore we can be confident our theory will do well in the neutral case, provided K is large and the diffusion constraints on r_0 and M are not violated.

The Markov chain is easily simulated as follows: each generation, a Poisson distribution with mean θ_i, where i is the current population size, is sampled to produce the size of the next generation. This continues until a steady state distribution is reached. (To check stationarity, we compute the relative error between stationary distributions generated from longer run lengths (total number of generations); if the relative error is below a certain error threshold, say 10^{-6}, we assume stationarity has occurred).

A Modified Wright-Fisher Model (MWFM)

In the Markov chain model, selection was not included and the island population was subject only to migration from the fixed mainland and density-dependent

regulation to ensure the population does not grow indefinitely. We wish to simulate a discrete version (i.e. non-overlapping generations) of the full continuous stochastic mainland-island model: individuals on the island carry a trait made up of many biallelic loci that is now subject to selection (one allele being selectively advantageous over the other on the island), and migrants, carrying a fixed number of deleterious alleles join the island every generation (see §4.2 and §2.1.1). The island population is still subject to density-dependent regulation. To do this, we simulate a modified version of the Wright-Fisher model. In the original Wright-Fisher model, the population size is kept constant, individuals mate randomly to produce the next generation and receive migrants from the a fixed mainland source (see §2.2.2). We modify the Wright-Fisher model to include growth and density-dependence regulation so that the population size is no longer constant and varies randomly. Below we describe the We deal with diploids here, but the model can easily be applied to haploids. The model is similar for the single locus and multilocus case except for an important difference between the average Wrightian fitness of the population, $\overline{W}_g$, and the average Wrightian fitness of a locus, $\overline{W}$, which we discuss below.

Consider a population with constant (effective) size N with (a single) allele frequency $p = \frac{n_+}{2N}$, where n_+ is the number of favoured alleles. In the classic Wright-Fisher model, if the population is subject to genetic drift only, the current generation is binomially sampled (with replacement) to produce the next generation. The number of favoured alleles in the next generation is $n'_+ = \text{Binom}(2N, p)$ and the allele frequency is $p' = \frac{n'_+}{2N}$, where $\text{Binom}(N, p)$ is a Binomial random number. Including selection and migration modifies p before sampling takes place (Ewens 1979).

In Fig. 4.19, the flow diagram of one life cycle in the MFWM is shown. The first half of each cycle deals with demographic changes, the second half with genetic changes.

Demographic stochasticity

Demographic stochasticity is included by Poisson sampling the local and migrant populations. Growth, (hard) selection and density-regulation are taken into account in one fell swoop to produce $N'' = \text{Poisson}(\theta_N)$ offspring where $\theta_N = N\overline{W}_g - \frac{r_0}{K}N^2$. Note that we have not assumed anything about how many loci contribute to fitness. To the offspring population, $I = \text{Poisson}(M)$ are added and the final size of the next generation is $N' = N'' + I$.

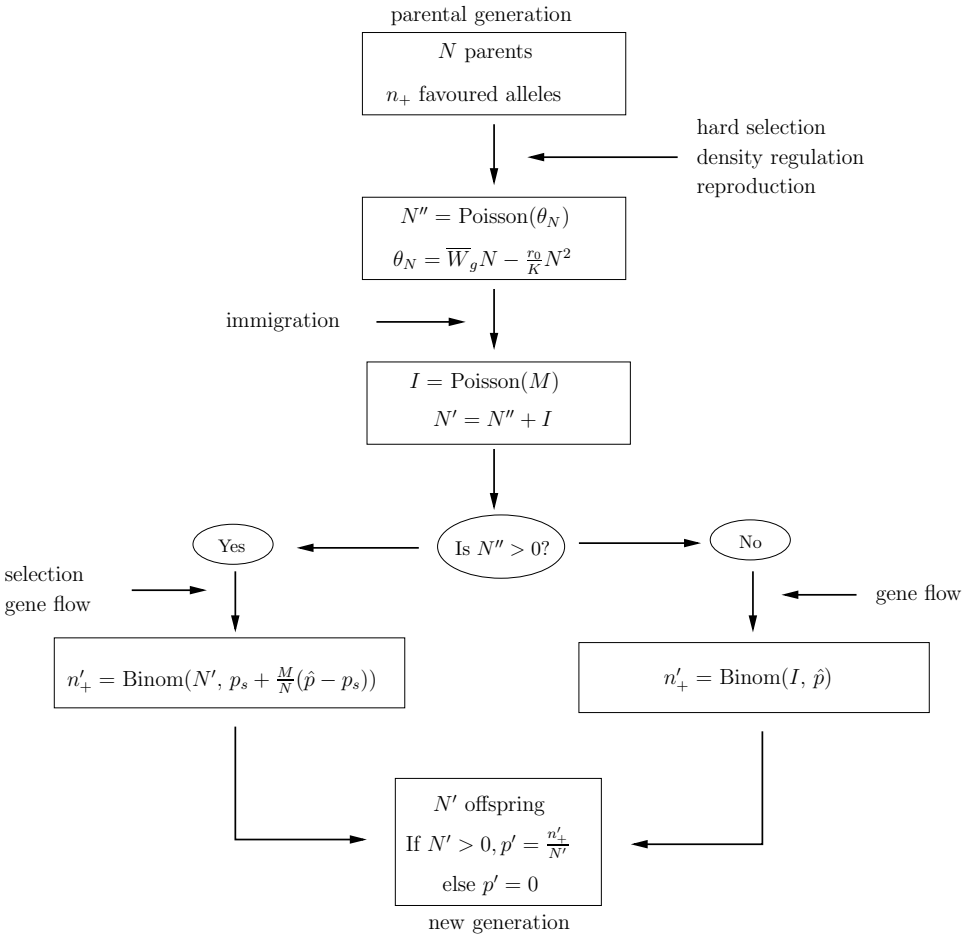


Figure 4.19 *Flow diagram of one life cycle in the Modified Wright-Fisher model.*

Genetic evolution

Suppose a single locus controls fitness. The allele frequency in the next generation, p', is found as follows. If there is at least one offspring ($N'' > 0$), then n'_+ alleles are generated by selection and migration. Selection acts on the local population and parents are chosen by their fitness. The allele frequency in the parental pool is $p_s = p\Big(p^2 W_{++} + pqW_{+-}\Big)/\overline{W}_1$, where W_{ij} is the Wrightian fitnesses of genotype ij and $\overline{W} = p^2 W_{++} + 2pqW_{+-} + q^2 W_{--}$ is the average fitness of the locus ($q = 1 - p$). During migration, gene flow 'replaces' a selected '+' allele by a migrant '+' allele at a rate $m\hat{p} = \frac{I}{N}\hat{p}$, and by a migrant '−' allele at a complementary rate $\frac{I}{N}\hat{q}$. Combining both processes, there are $n'_+ = \text{Binom}(2N', p'')$ favoured allele in the next generation, where $p'' = \frac{I}{N}(\hat{p} - p_s) + p_s$.

If no offspring are produced by the local population, n'_+ will depend solely on the migrant population, $n'_+ = \text{Binom}(2I, \hat{q})$. If $N' = 0$, the population goes

extinct in the next generation and of course $n'_+ = 0$. Finally, the allele frequency in the next generation is $p' = \frac{n'_+}{N'}$ if $N' > 0$ and 0 otherwise.

Multiple loci and the average growth rate

We define *two* fitnesses per genotype, one that contributes to the dynamics of p, $\overline{W}_{ij}$, and one that contributes to the average growth rate, $\overline{r}$, $\overline{W}'_{ij}$. The former is just the usual (Wrightian) genotypic fitness defined in §3.3. These determine the rate at which selection acts on the allele frequencies. The latter is defined so that the intrinsic growth rate in the MWFM $(1 + r_0)$ is correctly scaled by the number of loci: the W'_{ij} are $(\frac{1}{n}(1 + r_0) + s) : (\frac{1}{n}(1 + r_0)) : (\frac{1}{n}(1 + r_0) - s)$ for genotypes $++ : +- : --$ respectively. This guarantees that, on average, the average growth rate is $\overline{W}_g \approx 1 + r_0 + nsη(p - q)$ for n loci, which agrees with the continuous model. Note that for a single locus, the two types of genotypic fitnesses are the same (i.e. there is no need to scale r_0).

Simulation results

The simulation results from the Markov chain and MWFM have been presented throughout the thesis, in comparison with predictions from the continuous theory.

For the neutral results, we simulated the MWFM setting selection to zero. The neutral demographic results were checked against results from the Markov chain. The results were indistinguishable between these discrete models, as expected (see figures 4.2c, 4.2d and 4.5).

To check the nonneutral model we only simulated the MWFM. Overall, we found that the continuous model is qualitatively, and in some cases quantitatively, a reasonable approximation of the discrete Modified Wright-Fisher model. This is not surprising given the MWFM satisfied many of the restrictive assumptions upon which the continuous theory is based, such as linkage equilibrium. The approximation becomes more accurate the greater the intrinsic carrying capacity K; surprisingly it does well for seemingly small values too, e.g. $K = 10$, as the MET and fixation probability results show. This is because even for a small carrying capacity the population can attain sizes many orders of magnitude higher.

Nevertheless, there are several issues. First, though the average stationary population size and allele frequency are in good agreement for the parameter values shown, the continuous and discrete variance are not. For example, in Fig. 4.9, the discrete variance σ^2_N is higher for $K = 20$ than the continuous theory predicts. Similarly, for $\hat{p} = 0.01$, in Fig. 4.5, both demographic and

genetic variances are in poor agreement with the continuous theory for $K = 20$. This is because for the migration rates shown, the continuous theory violates the diffusion constraint: the fraction of migrants entering the population, relative to the intrinsic carrying capacity, is too high. Recall that migration does not contribute to fluctuations directly in the continuous dynamics, whereas in the discrete model this is approximately the case in the limit of large K and $M \sim O(1)$. Furthermore, allele frequencies in the discrete models can change typically at most by $\sim \frac{1}{K}$. Therefore if $\hat{p} < \frac{1}{K}$, migrant alleles will not be adequately represented. Hence K is chosen so that this is not the case. Note however that in general we expect the diffusion approximation to break down when there are few copies of the local allele in the population or when there are few individuals (Ewens 1979).

Second, the continuous theory consistently predicts a higher migration rate M_{max}, the rate at which the genetic and demographic variance is a maximum. This can be seen in the 'shifted' peak probabilities, figures 4.4 and 4.15, and also in the marginal distributions, e.g. Fig. 4.14. In the latter, the maladaptive peak is consistently higher than the theoretical distribution. This suggests that in the MWFM the population does not visit the high state peak as often as the continuous dynamics predict, even for a population with an intrinsic carrying capacity as large as $K = 200$. This is most likely because in the discrete model migration contributes to demographic stochasticity, whereas it does not in the continuous model.

Third, the single locus transition rates and mean exit times were orders of magnitude off the continuous predictions. This is perhaps due to the assumptions the theoretical calculations were based upon, rather than failure of the continuous theory as a whole, especially since the discrepancy between the continuous and discrete ratios R were in better agreement. This requires further investigation.

Finally, we have used large values for the unscaled selection coefficient in our theoretical predictions. However, in the discrete simulations, the unscaled values are used, and these large values are scaled by r_0. Therefore, provided r_0 is small, the unscaled selection coefficients will be small. We discuss this issue in more detail in §4.6.

We next consider how genetic and stochastic drift affect phenotypic evolution under the assumption of constant variance.

4.5 Phenotypic evolution

In the last chapter, we investigated a deterministic quantitative trait model with population dynamics, see §3.5. The crucial assumption of constant variance simplified the trait dynamics and resulted in a qualitatively different picture of local adaptation under maladapted gene flow from the explicit genetic models.

Here we build upon this model to investigate the effect that genetic and demographic drift have on trait evolution under directional selection. We can expect the variance to be reduced by genetic drift at a rate inversely proportional to the effective population size (Crow & Kimura 1970). Demographic stochasticity will further reduce the additive genetic variance. If migration is high enough to prevent frequent extinction, this may counteract the effect of demographic stochasticity, inflating the genetic variance. However, high migration rates will introduce positive linkage disequilibrium into the population. As we will see however, quite low values of migration and selection significantly reduce the risk of extinction at stochastic equilibrium; weak migration and selection relative to high recombination rates will break down nonrandom associations between loci.

Next we derive the joint stationary distribution of trait mean and population size. We will find that the stochastic model closely mirrors the deterministic behaviour, as expected when drift is weak. For certain parameter values however, the stationary distribution is unbounded.

4.5.1 Stochastic equilibrium

To derive the stationary distribution for population size and the trait mean, $\Psi_{N,z}$, we follow the usual route of solving a bivariate forward Fokker-Planck equation with the appropriate infinitesimal moments (Gardiner 1985).

Recall the deterministic equation for the trait mean z assuming a constant variance v_z, (cf. Eq. 3.22). The stochastic dynamics are

$$dN = \Big(\overline{r}N - N^2 + M\Big)dt + \sqrt{\frac{N}{\zeta}}dW_N, \tag{4.51a}$$

$$dz = \Big(v_z\partial_z\overline{r} + \frac{M}{N}(\hat{z} - z)\Big)dt + \sqrt{\frac{v_z}{N\zeta}}dW_z. \tag{4.51b}$$

As in the genetic stochastic model, the bracketed terms multiplying 'dt' are just

the deterministic dynamics, see Eq. 3.22; the square-root terms represent the stochastic contribution to the dynamics. Note that this term in Eq. 4.51b has the same origin as the stochastic term in Eq. 4.1b. The infinitesimal moments for the FFPE equation are therefore

$$A_N = N\partial_N U, \tag{4.52}$$

$$A_z = \frac{v_z}{N}\partial_z U, \tag{4.53}$$

$$B_N = \frac{N}{\zeta}, \tag{4.54}$$

$$B_z = \frac{v_z}{\zeta N}, \tag{4.55}$$

where potential U is

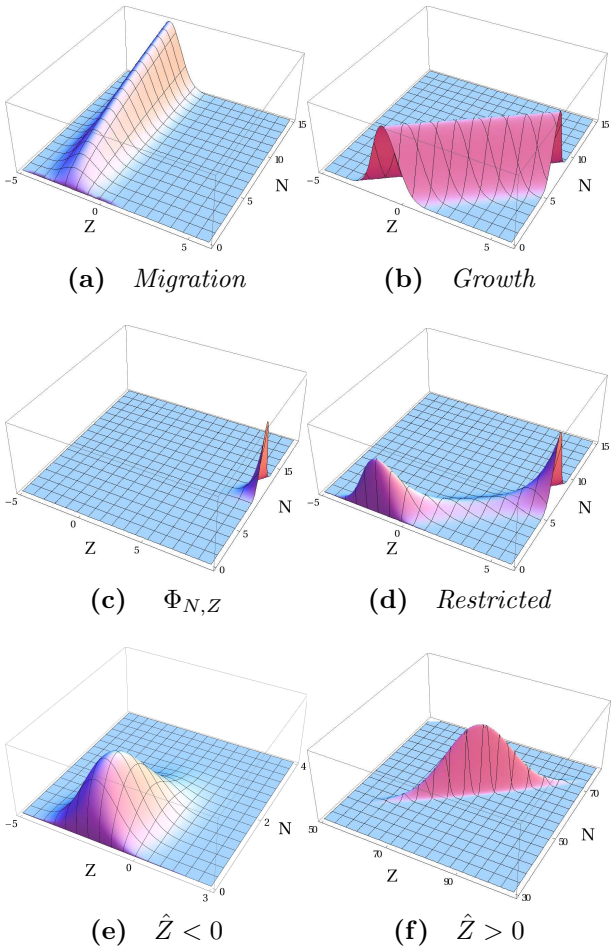
$$U = \overline{r}N - \frac{1}{2}N^2 + M\ln N + \frac{M}{2v_z}(\hat{z} - z)^2 \tag{4.56}$$

Rescaling the variables as we did in the last chapter, $S = s\sqrt{v_z}$, $Z = \frac{z}{\sqrt{v_z}}$ and $\hat{Z} = \frac{\hat{z}}{\sqrt{v_z}}$, and solving the homogeneous FFPE Eq. 4.2, the stationary distribution is

$$\begin{aligned}\Psi_{N,Z} &= \frac{1}{Z_0}\frac{e^{2U}}{N} \\ &= \frac{1}{Z_0}\Big(N^{2M\zeta-1}e^{-M\zeta(\hat{Z}-Z)^2}\Big)\Big(e^{\zeta(N-\overline{r})^2}\Big)\Big(e^{\zeta(\overline{r})^2}\Big).\end{aligned} \tag{4.57}$$

where as usual Z_0 is the normalisation constant. The distribution is determined by three different terms. The contribution of each term to $\Psi_{N,Z}$ is shown in Fig. 4.20 for parameter values $S = 1$, $M = 0.8$ and $\hat{Z} = -2$.

The first bracketed term in Eq. 4.57 is due to migration. If migration dominates, the Gaussian factor pulls the stationary distribution towards the average of the migrant population, $\hat{Z}$. In the $N - Z$ plane, these terms produce a line of Gaussians centered about $\hat{Z}$, see Fig. 4.20a. The height of the Gaussians are modified by the demographic factor multiplying the exponential; it will dominate when N is small and migration is low, i.e. $M < \frac{1}{2}$. The second term is also a demographic Gaussian, but due to growth, density regulation and hard selection. This dominates when selection is stronger than migration and drift, pulling the distribution towards the average fitness of the population. It produces a line of Gaussians, each one centered about a point on the line $1 + sz$, see Fig. 4.20b. The



(a) *Migration* **(b)** *Growth*

(c) $\Phi_{N,Z}$ **(d)** *Restricted*

(e) $\hat{Z} < 0$ **(f)** $\hat{Z} > 0$

Figure 4.20 *Phenotypic distribution* $\Phi_{N,Z}$ *(unnormalised) for* $S = 1$, $M = 0.8$ *and* $\hat{Z} = -2$, *unless otherwise stated. (a) Migration term, (b) growth term and (c) distribution* $\Psi_{N,Z}$, *in which the selection term dominates. (d)* $\Phi_{N,Z}$ *appears bimodal if we restrict the range of the distribution. The interior peak maximum is approximately at the equilibrium state* $\{1.106, -0.618\}$. *Examples of* $\Phi_{N,Z}$ *when: (e) migration dominates selection,* $S = 0.4$, $M = 0.7$, $\hat{Z} = -2$ *with corresponding deterministic equilibrium* $\{1.091, -1.377\}$, *(f) the local allele is common:* $\hat{Z} = 0.2$, $S = 0.7$ *and* $M = 0.5$, *with corresponding deterministic state* $\{57.435, 80.609\}$.

final term is a pure exponential due to selection on the trait mean. It will easily be the dominating factor when migration is weak relative to selection. This is evident from Fig. 4.20c, which shows the product of all three terms, giving the unnormalised stationary distribution, $\Phi_{N,Z}$.

If we zoom in to a smaller range of N and Z, the small, interior peak centered approximately about the deterministic state is visible, see Fig. 4.20d. This distribution is unbounded and cannot be normalised, since the exponential term is dominant. This happens under maladaptive gene flow, when selection is

stronger than migration i.e. when the migration rate is below critical rate M_{div}, the rate at which the deterministic equilibria diverge and are indeterminate; see §3.5.2.

For migration rates above M_{div}, the distribution is bounded and centers about the corresponding deterministic equilibrium state, see Fig. 4.20e. This is similar to the maladaptive peak seen in the genetic models: the trait mean fluctuates about a narrow range of values, which may or may not lead to negative growth rates. In contrast to the genetic models, the population can be maladapted, and still attain population sizes several times the intrinsic carrying capacity.

If the favoured allele is the more common allele on the mainland, $\hat{Z} > 0$, the distribution is centered about the deterministic state, which has a large population size and high trait mean, see Fig. 4.20f. The domain of the peak represents the population's ecological niche. Notice that only a narrow range of states are probable however.

The stochastic trait model is a viable alternative to the genetic model only when migration dominates selection under maladaptive gene flow, or when the favoured allele is common on the mainland. Otherwise the stationary distribution is unbounded, due to exponential selection. The fact that the phenotypic distribution is never bimodal can be reconciled with the fact that bimodality in the multilocus stationary distribution is fleeting when loci numbers are large, which is a key assumption for the phenotypic model. However, in the latter this marks a transition between adaptation and maladaptation, whereas in the former there is no such transition. In this limited sense then, both types of models give a qualitatively similar picture of local adaptation under maladapted gene flow and stochastic effects.

4.6 Summary and discussion

In this chapter we have explored the central theme of the thesis, namely the effect of genetic and demographic drift on local adaptation. We have investigated two different continuous models based on the deterministic models of the last chapters: a multilocus genetic model in which n independently evolving loci contribute additively to trait, and its phenotypic equivalent, in which only the trait mean evolves alongside population size. We have explored the stationary properties of both models, and properties of the stochastic trajectories in the former. Continuous models can be treated as diffusion approximations of an

infinite number of discrete models. Since we focussed mostly on the genetic model, we also tested its validity as an approximation of discrete models by comparing it to a modified version of Wright-Fisher model, in which population sizes can fluctuate.

Adaptive and maladaptive ecological zones: the basic observation that local adaptation is possible, albeit at the cost of a migration load, when selection is strong relative to migration remains true, provided drift is weak. Conversely, high migration will result in maladaptation. Demographic stochasticity produces a third possibility however. The genetic model predicts that for a certain range of parameter values, when the favoured allele is less common on the mainland than the deleterious allele, the population can shift between two 'ecological zones': an adaptive zone and a maladaptive zone. The adaptive zone is the population's ecological niche: the favoured allele is established at high frequencies and the population is many times the intrinsic carrying capacity in size. Within the maladaptive zone, the favoured allele cannot spread and the population becomes a pseudo-sink (Pulliam 1988). The population is particularly vulnerable in the maladaptive zone because not only is it genetically depauperate, which can limit future adaptability (Willi et al. 2006), small population numbers may lead to extinction from inbreeding depression (O'Grady et al. 2006) or from Allee effects (Keitt et al. 2000), as discussed in the previous chapter, §3.6.

Bimodal range: we found that transitions between states was possible for a limited range of migration rates, the 'bimodal range'. The population is less likely to be found within its ecological niche as migration increases across this range i.e. the ecological niche becomes less dominant and vice versa for the maladaptive zone. At migration rate M_{max}, defined to be the center of the bimodal range, genetic and demographic variation are maximal. Since genetic variation is necessary for future adaptability, the population has the best chance of adapting to a sudden change in the local environment at rate M_{max}. We demonstrated the basic dependence of the bimodal range on parameters s, n, $\hat{p}$ and ζ. Analogous to the deterministic model from the last chapter, we found that as the number of loci n increases, bimodality exists for weaker selection and migration. Crucially, for a given selective strength, the bimodal range becomes narrow as the number of loci increases; the bimodal range becomes a sharp threshold above which adaptation is not possible and bimodality disappears. This suggests that the genetic model behaves more like soft selection models in the limit of large n. Clearly the number of loci matter. In contrast Holt et al. (2003) found that for $n \geq 5$, the number of loci do not affect the results. This can be

reconciled with our results, since for a large number of loci, bimodality will be fleeting, irrespective of the number of loci. But loci numbers will still affect where this threshold lies, as in the deterministic models.

Transitions between zones: drift provides a means of escape from maladaptation. The accumulation of small fluctuations can shift the population out of a maladaptive state and into a niche state. This can be likened to niche evolution (Holt 1997, Holt & Gomulkiewicz 2002, Kawecki 2000, Holt 2009). We determined an asymptotic upper bound for the rate of shifts between zones using the mean time it takes for the population to exit a zone. We found that the shift rates are exponentially distributed across the bimodal range, as expected; transitions to the ecological niche become less likely as migration increases. The depth of the potential well (i.e. the barrier that must be surmounted by the population) is the dominant factor in the rate transitions, as it is for many rate calculations (see references in §4.3.4). Our calculations are in qualitative agreement with results obtained from simulations. It would be interesting to compare the discrete results with more accurate shift rates. One approach would be to determine the inverse of the dominant eigenvalue of the backward Fokker-Planck operator Eq. 4.2, as was carried out by Barton & Rouhani (1987*a*) for a soft selection model of multilocus evolution. Since many traits are polygenic (Bulmer 1980) and under weak directional selection (Kingslover et al. 2001), transitions between zones may be unlikely to occur in real populations under maladaptive gene flow.

When drift is strong: demographic drift causes the population to go extinct. We have used the discrete models to investigate the mean time until extinction in the case of a single locus. Simple intuition would suggest that, in general, the mean extinction time should increase the larger the carrying capacity and/or the intrinsic growth rate. We found, however, that the initial genetic composition affects the mean extinction time so that it may be reduced should the initial allele frequency be low enough.

It has been argued that the mean extinction time is not a good quantifier of extinction risk, and that the probability of extinction is better (Ludwig 1996). Furthermore, use of diffusion to approximate these quantities has been criticised (Grasman & Ludwig 1983). Nevertheless, the mean extinction time remains an important quantity in conservation studies (see Ch. 2, §**??**) and as our simulations have shown, diffusion approximations can be accurate for populations with small intrinsic carrying capacities, e.g. $K = 10$. Furthermore, empirical studies show that genetic factors impact upon extinction risk and cannot be ignored (O'Grady

et al. 2006, Spielman et al. 2004). We therefore believe it would be fruitful to develop theoretical results to support our simulations. The most obvious way to do this would be to solve the appropriate inhomogeneous two-dimensional backward Fokker-Planck equation (see Eq. B-17, App. B).

The diffusion approximation: to assess the validity of the genetic model, we compared it to a discrete model, the Modified Wright-Fisher model. We found that for the parameter values explored, the continuous theory was in good qualitative agreement with the stationary properties of the MWFM for weak to moderate drift, moderate to large carrying capacities, $K \geq 20$, and small migration rates, relative to the carrying capacity, $M/K \ll 1$. (Failure of the continuous model to correctly predict the transition rates in the single locus case has been discussed in §4.4.4). This is unsurprising given the discrete model satisfies many of the restrictive assumptions underlying the continuous model, such as the guarantee of linkage equilibrium. The analytical approximations we developed for the continuous theory also agree reasonably well with the discrete model, provided drift is moderate to weak. However, if the local allele is sufficiently rare or drift is only moderately weak, 'ghost peaks' may form at the boundaries of the stationary distribution. Then, bimodality has no deterministic counterpart and our approximations cannot be used.

We expect the diffusion approximation to break down once the expectation of the third (or higher) raw moment of generational changes in N and p_j are non-zero, see Eq. 4.48 (Ewens 1979, Karlin & Taylor 1975*b*). In the single locus case, we found that for seemingly large values of the scaled parameters, e.g. $s = 8$ and $M = 17$, the diffusion approximation holds. This is because despite the large values of the scaled parameters, the simulations were run using the unscaled parameters, and these did not violate the diffusion constraints given in §4.4.4.

Realistic parameter values: it has not been our aim to develop models that can be directly related to empirical studies. Nevertheless, we wish to say something about the order of magnitude of the parameters and whether they fall within realistic bounds. Unfortunately this question has no straight-forward answer, not least because reliable estimates of genetic and ecological parameters are hard to obtain. Also data on the genetic basis of traits is sparse. However, we can use the empirical estimates of selective strength and gene flow for the water snake *Nerodia sipedon* (King & Lawson 1995), which shows variation in banding at a single locus (Hedrick 2006), to obtain a crude order-of-magnitude estimate of the parameters. The water snake lives in Lake Erie on several islands. The unbanded form of the snake is commonly thought to be an adaptation for

camouflage against the limestone background of the island this form is commonly found on. The mainland consists of the banded forms. The selective advantage of the unbanded form was found to lie between $s = 0.11$ and $s = 0.28$ (King & Lawson 1995). Gene flow was estimated at $m = 0.01$. Suppose $Kr_0 = 1$ and that $M_u = M = mK$. In the single-locus simulations we used $M \approx 17$. To reproduce such a high migration rate, given gene flow rate m, the intrinsic carrying capacity would be $K = 1700$. We can use the estimate for K to determine the unscaled selection coefficient, since $s_u = r_0 s$, where $r_0 = 1/K$. This gives $s_u \approx 0.005$, which is too small for the water snake. It is orders of magnitude smaller than the estimate $s = 0.16$ averaged across species (Kingslover et al. 2001). Where the single locus model may fail to use parameters within realistic ranges, the multilocus model will fare better, though there may be other issues, such as linkage disequilibria and epistatic interactions. More worryingly, however, is the estimate for r_0. Empirical data on the intrinsic growth rate is confused by different definitions and empirical methods, but is typically greater than $r_0 \geq 0.02$ (e.g. see Fagane et al. 2010). This suggests that the diffusion model requires smaller growth rates than are likely to be found in nature.

Trait evolution: we found that the quantitative trait model for directional selection, assuming the additive genetic variance is constant, behaved markedly differently from the genetic model. Under maladaptive gene flow, the model cannot adequately describe local adaptation: if selection exceeds migration, the stationary distribution is unbounded. For large migration rates that exceed M_{div}, the stationary distribution is bounded and the population will most likely be maladapted, though the population size may be large due to migration, depending on the average migrant trait. In other words, adaptation is severely limited under demographic and genetic drift, if we make the usual assumption of constant phenotypic variance. Clearly, assuming the genetic variance is constant is inappropriate and this assumption must be made with caution.

Comparison with previous studies: simulation studies of stabilising selection under genetic and demographic drift have been studied by Holt et al. (2003). Our results are qualitatively similar to theirs, in that they also observed bimodality and determined a probability of adaptation similar to our peak probability $P(+)$. This is because free recombination between loci was also assumed in their model.

Issues and future work: genetic drift causes random associations between loci, which may or may not attenuate the effect of selection. Again, high recombination will break down associations between loci. However, little

is understood about the build up of nonrandom associations under both demographic and genetic drift, and density regulation. As is clear from the quantitative trait model, assumptions made about the underlying genetics and associations between loci can greatly alter the results. We therefore believe it would be fruitful to investigate linkage disequilibrium in, say, a two-locus model under genetic and demographic drift.

The order of selection and migration are unimportant in the continuous model (this is imposed by the diffusion constraints). For more realistic discrete models, however, this is not the case (Holt et al. 2003). Furthermore, the assumption of freely recombining loci does not take into account proper sexual reproduction. More light could be shed on both the issue of life history events and the effect of sex (and therefore nonrandom associations) with an individual-based model that incorporates proper sex into the eco-evolutionary dynamics.

Despite its shortcomings, we believe the continuous description framework is a valuable one. For example, it would be relatively straight-forward to investigate other models of selection, such as stabilising or disruptive selection. The assumption of a constant environment was necessary as a first step towards understanding the effect of stochastic events on local adaptation under eco-evolutionary dynamics. However, future work could involve relaxation of this assumption, in addition to all the others mentioned above.

Next, I generalise the mainland-island model to the infinite island model and examine the conditions for local adaptation under deterministic and stochastic local dynamics.

Chapter 5

The island model

In this chapter I extend the mainland-island model to a metapopulation of an infinite number of islands. Each island is loosely coupled to all others by migration and evolves according to either the deterministic or stochastic dynamics of the previous chapters. In the case of two alleles per locus, the metapopulation reduces to two habitats: one habitat, a fraction ρ of the metapopulation, favours the '+' allele, the other habitat, of size $1-\rho$, favours the '−' allele. Islands within a given habitat experience identical local dynamics, and each habitat behaves as one contiguous island. Therefore the metapopulation effectively behaves as a two-island model in which there is migration between the islands, unlike the mainland-island model. With local stochastic dynamics, the assumption of an infinite island number of islands simplifies analysis and ensures the distribution is a smoothly varying function across demes.

I show that the infinite island model under deterministic local dynamics behaves similarly to the two habitat models discussed in §2.3.1; under certain conditions either specialists or generalists evolve. I also show that the inclusion of multilocus dynamics can severely constrain the range of conditions for which adaptation within a given habitat is possible. If the metapopulation suffers a catastrophe, this can invert any source-sink structure that may emerge through the coupled dynamics of population size and allele frequency. Finally, I show that stochastic effects can reduce demic sizes and increase mixing between habitats with increasing isolation. This, however, comes at the costs of low average population sizes, increasing the risk of extinction. Before introducing the model, I review previous work.

5.1 Previous relevant work

Wright's infinite island model has been modelled in a heterogeneous environment (see Ch. 2). Under deterministic dynamics however, Wright's island model in a heterogeneous environment collapses to the two habitat models discussed in §2.2.

Barton & Whitlock (1997) found that in an infinite island model with fixed deme size N_e, an allele with selective advantage s in a fraction ρ of the entire population and a disadvantage $-\gamma s$ in the remainder, can be established below a critical rate $m_L = \gamma s/(\gamma(1-\rho)-\rho)$. Alternatively, the allele can invade the island population for fractional habitat sizes greater than $\rho_L = (1 - s/m)/(1 + 1/\gamma)$, where $m = M/N_e$. A similar condition in terms of the ratio of selective differences, γ, was first derived by Bulmer (1972) for two habitats. The allele is always established if $m < s$. This is because selection only varies in direction and not in strength between the two habitats. The alternative allele is established if ρ is at most as large as a critical fraction ρ_H. Barton and Whitlock also considered the effect of genetic drift and extinctions on establishment of the allele. If extinctions occur at a rate λ, with recolonization occurring immediately after extinction events and relatively weak migration ($m \ll s$), the critical fraction becomes $\rho_L = (1 - s/(m + \lambda))/(1 + 1/\gamma)$ in the deterministic limit. Migration and extinction increase the rate of gene flow, having a similar effect on local adaptation. Again, the allele can always be established provided selection dominates ($s > (\lambda + m)$). They found that polymorphism becomes less likely as the extinction rate increases. Furthermore, if drift is strong (i.e. $N_e m \ll 1$) and the two habitats are roughly equal, $\rho \approx 0.5$, polymorphism cannot be established in the metapopulation; depending on initial conditions, one of the two alleles is fixed.

Ronce & Kirkpatrick (2001) study the evolution of a quantitative trait in a model of two discrete habitats with explicit population dynamics including logistic regulation. Habitat heterogeneity is set by the relative difference between the optimum average trait of each habitat. They find that for high habitat heterogeneity (large variation in selection or habitat quality) and intermediate to high migration rates a source-sink structure emerges: the source habitat has a high population density and will be close to perfect adaptation, whereas the maladapted sink habitat depends on migration from the source. Since most individuals are found in the source habitat, a specialist species evolves. If migration is too high, it is possible for "migrational meltdown" to occur and the sink habitat goes extinct. With less heterogeneity, habitats become equally

populated with identical trait averages, forming a generalist species. Both equilibria are stable for very small migration rates and high habitat heterogeneity. Furthermore, the authors compared a discrete version of their trait model with an explicit multilocus model using simulations and found the qualitative behaviour of both models to be in agreement.

Filin et al. (2008) extend this model and investigate the effect that different forms of density regulation has on the evolution of generalists and specialists. They find that the specific form and strength of density regulation strongly influences the conditions for a generalist or specialist outcome; increasingly concave density regulation functions (stronger dependence at high densities, weaker dependence at low densities) improves conditions for a generalist species. This is similar to the equilibrium results of Kawecki & Holt (2002), Spichtig & Kawecki (2004) models.

Barton & Rouhani (1993*a*,*b*) examine an infinite island model, in which a quantitative trait evolves under stabilising selection, migration and genetic drift with soft (Barton & Rouhani 1993*a*) and hard (Barton & Rouhani 1993*b*) disruptive selection, (in the latter, changes in population size are implicitly accounted for). They investigate the third phase of Wright's "shifting balance", which concerns the spread of a fitter combination of genes from one deme to all others. Wright's "shifting balance" (Wright 1931, 1932, Coyne et al. 1997, Peck et al. 1998, Turelli et al. 2001, Singh & Morton 2004) was postulated for structured populations in a homogeneous environment (uniform selection across habitats). Although they study a different mechanism for evolution, the method they employ to analyse the infinite island model is applicable to the infinite island models I consider.

The infinite island model I examine in this chapter is an extension of the mainland-island work of the last two chapters. It also bridges the gap between the above studies by including habitat size (fraction or frequency) ρ, genetic dynamics and genetic drift. Furthermore, I consider genetic and demographic stochasticity as well as explicit multilocus dynamics, which is novel. To treat the stochastic case, we follow the method used by Barton & Rouhani (1993*b*,*a*). I will show that the under single-locus dynamics, the deterministic island model agrees with soft selection models (Levene 1953, Maynard Smith 1966) in that increased isolation promotes polymorphism. I also show that, as expected, an allele can invade either if its habitat is more frequent or if it has a greater (absolute) selective advantage. When habitats are equally frequent, the 'hard demographic' advantage of a larger

habitat size disappears. In this case, the species evolves into either a generalist or specialist species', depending on the initial state of each habitat and the relative strength of migration and selection. Crucially, under (symmetric) multilocus dynamics, the conditions for genetic polymorphism are greatly reduced and the species switches between specialist states as migration increases. If one habitat suffers a catastrophe, the source-sink structure that emerges when one habitat is rare can be reversed. Finally, I show that demographic and genetic drift can reduce average demic sizes. At small migration rates, if selection is weak and drift is strong, habitats become more mixed on average despite increased isolation. This could lead to extinction of the entire metapopulation due to severe migration loads in each habitat, high demographic stochasticity and reduced colonization.

The rest of the chapter is set out as follows. In Sec. 5.2 I define the general stochastic model for multiple loci. I then solve the deterministic mode for a single locus, first for varying habitat frequency and a fixed migration constant, and then vice versa, §5.3.1. I also calculate the habitat range of frequencies for which genetic polymorphism is guaranteed. Next, I consider the effect of a catastrophe on the population. If the catastrophe is severe enough, the sink population has a chance of evolving. After examining the multilocus case, I move on to examine the stochastic model in Sec. 5.4.1. In Sec. 5.5, I summarise and discuss the results.

5.2 Model and dynamics

The island model consists of an infinite number demes that are connected by migration (which we discuss in detail below). Each deme experiences similar conditions to the island population of the last two chapters. On each island, diploid (we ignore haploidy) individuals carry a trait expressed by one or many biallelic loci ('+'/'−' alleles). Selection (specifically directional selection i.e. fitness is an exponential function of trait) acts on the trait. The '+' allele is favoured in some demes, whereas the '−' allele is favoured in others; this leads to our definition of a habitat, which we discuss in more detail a bit later. If we assume that loci evolve independently of each other, that is, nonrandom associations between loci are broken down at a much faster rate than selection, then we can follow the dynamics of individual loci, just as in the genetic mainland-island models of the last two chapters. In this chapter, we focus exclusively on the explicit genetic dynamics of individual loci and do not investigate the phenotypic model, which follows the average trait dynamics instead (see §3.5 and §4.5). To

ensure population numbers are regulated on each island, we once again include density-dependent regulation in the local dynamics. The stochastic effects of finite-sized populations can be taken into account in much the same way they were in Ch. 4.

Migration

The key difference between this model and the mainland-island model of the last two chapters is migration. In the mainland-island model, migration was unidirectional (from mainland to island only) and occurred at a fixed rate M. Furthermore, the migrant pool was assumed to be so large that its average size and genetic structure could be treated as fixed and constant in time. In this model we relax both these assumptions.

Every generation, each deme receives migrants from a common migrant pool that is randomly sourced from all demes without preference. Since the migrant pool is a property of all demes, it evolves alongside the metapopulation until stochastic equilibrium is reached. Let $\hat{x}$ be the average of x in the migrant pool at a given instant. Following Barton & Rouhani (1993*a*,*b*), averages taken across the metapopulation, denoted $\langle x \rangle$, determine the (instantaneous) state of the migrant pool, $\langle x \rangle = \hat{x}$. As in their work, we distinguish between these two averages: metapopulation averages $\langle x \rangle$ determine the migrant pool, which in turn has average properties $\hat{x}$. The rate of migration into a deme is therefore proportional to the average size of the migrant pool, $M\hat{N}$, where M is a fixed constant that tunes the strength of migration relative to selection and random drift.

In contrast to the mainland-island model, each island sends out emigrants at a rate proportional to its size: the larger a deme is, the greater the number of individuals it sends out to all other demes and hence the greater its influence on other demes via the migrant pool.

Habitat definition and dynamics

Selection can vary in direction and in strength from deme to deme i.e. one allele is advantageous and disadvantageous in different demes. We define a habitat i as the fraction of demes ρ_i that experience the same selection pressure for the same allele.

Within a habitat, each deme evolves according to the *same* local dynamics. The most general dynamics include stochastic effects due to random births and deaths on each island. Under local stochastic dynamics, the demic population sizes and allele frequencies within a habitat are distributed about averages taken

across the habitat. If the metapopulation is divided into L habitats, so that $\sum_{l=1}^{L} \rho_l = 1$, the contribution of habitat l to the migrant pool is weighted by ρ_i, $\hat{x} = \sum_{l=1}^{L} \rho_l \langle x_l \rangle$, where $\langle x_i \rangle$ is the average across habitat i.

For a deme within habitat i, we can write down the coupled stochastic dynamics of the population size, N_i, and the allele frequency (of the '+' allele) at locus k, p_{ik}, when selection, growth, migration and finite-size effects act upon the deme. These dynamics are similar to the coupled equations for N and p_j in the deterministic (see §3.2.1) and stochastic (see §4.2) mainland-island models, since the processes of selection, growth, immigration, gene flow and population regulation are similar. However, as explained earlier, the key difference in the dynamics is that an island in this model experiences emigration.

The dynamics can be written as stochastic differential equations, similar to the dynamics of the stochastic mainland-island model, see Eq. 4.1. Using the scaled variables (cf. §3.2.1), the stochastic dynamics are

$$dN_i = \Big(N_i(\overline{r}'_i - N_i) + M(\hat{N} - N_i) \Big) dt + \sqrt{\frac{N_i}{\zeta}} dW_{N_i}, \tag{5.1a}$$

$$dp_{ik} = \Big(p_{ik}(1 - p_{ik}) \partial_{p_{ik}} \overline{r}'_i + \frac{M(\widehat{(Np_k)} - \hat{N} p_{ik})}{N_i} \Big) dt + \sqrt{\frac{p_{ik}(1 - p_{ik})}{2\zeta N_i}} dW_{p_{ik}}, \tag{5.1b}$$

where $\overline{r}'_i = 1 + \beta_i \sum_k \alpha_{ik}(2p_{ik} - 1)$ is the average fitness of a deme, β_i is the selection pressure, α_{ik} is the contribution of the k^{th} locus on the trait, $\zeta = r_0 K$ is the scaling parameter (cf. 4.2), and r_0 and K are the intrinsic growth rate and maximum population size in the absence of selection and migration. Let us take a closer look at the dynamics. The first term in Eq. (5.1a), is due to logistic regulation (Murray 1989), as explained in §3.2.1: the population grows at a rate set by its average fitness, $\overline{r}'$ (this is the same as $\overline{r}$ from previous chapters, when many loci contribute to trait). To ensure the population does not get large indefinitely, the quadratic term in N_i imposes a penalty on the overall growth rate, $\overline{r}'_i - N_i$, as the population becomes too large. The second term accounts for the net movement of migrants between the deme and the rest of the metapopulation per generation; $M\hat{N}$ migrants come into a deme and MN_i migrants leave a deme. The final term is due to demographic stochasticity; random births and deaths in a finite population are modelled as a Poisson process (Karlin & Taylor 1975*a*) (see also §4.2).

The first term in Eq. (5.1b) accounts for selection acting on locus k (Ewens 1979). The unusual second term in Eq. (5.1a) is due to gene flow (note the ratio M/N_i) from all other demes. To see this, suppose the metapopulation consists of d demes, where d is finite. Using the relation $\langle x \rangle = \hat{x}$, gene flow into an island gives

$$\frac{1}{d}\frac{\sum_{j=1}^{d} N_j(p_{jk} - p_{ik})}{N_i} = \frac{(\widehat{(Np_{ik})} - \hat{N}p_{ik})}{N_i}. \tag{5.2}$$

In contrast to constant population size models and the models of the last two chapters, the (net) rate of gene flow depends on the difference in (mal)adaptedness between demes and the ratio of population sizes between pairs of demes. Abundant populations contribute more to the gene pools of smaller populations than the reverse (Kawecki & Holt 2002). Of course, migrants only contribute to gene pools that have adapted differently from their deme of origin (i.e. $p_j \neq p_i$).

We will consider only two habitats. Habitat 1 consists of a fraction $\rho_1 = \rho$ of demes in which the '+' allele has a selective advantage $s_1 = s$ ($s > 0$). The rest of the metapopulation is habitat 2, where the favoured allele is deleterious with selective disadvantage $s_2 = -\gamma s$ and $\gamma > 0$. An allele is considered to have invaded a habitat if it can be found there, even at low frequencies. If it is more frequent than the resident allele, it has 'successfully' invaded or 'spread' through the habitat. Of course, successful invasion is meant only in the genetic sense; demographic benefits are unlikely to follow genetic invasion. For simplicity, we set $\alpha_{ik} = 1$, so that $s_i = \beta_i$. We will use the terms 'frequency', 'size' and 'fraction' interchangeably to refer to ρ.

How large does ρ have to be to ensure the '+' allele can invade habitat 2? Alternatively, how weak does selection for the resident allele have to be for invasion to be possible? What effect does genetic invasion or genetic polymorphism have on the distribution of individuals between habitats i.e. do generalists or specialists evolve? Do stochastic fluctuations significantly affect the conditions for invasion? We will answer these questions as we explore the model.

We next investigate the deterministic model, focusing on a single locus. We examine two complementary cases: (i) habitat frequency ρ varies and migration constant M is kept fixed, and (ii) the converse, where ρ is fixed and M varies. Before moving on to the multilocus case, we investigate the effect of a catastrophe on the metapopulation.

5.3 Deterministic demes

In the absence of stochasticity, there is no longer a distribution of population sizes or allele frequencies within each habitat; the demes within a habitat are identical in size and genetic composition. The deterministic dynamics are,

$$\frac{dN_i}{dt} = N_i(\overline{r}_i - N_i) + M\rho_j N_j, \tag{5.3a}$$

$$\frac{dp_{ik}}{dt} = p_{ik}(1-p_{ik})s_i + \frac{M\rho_j N_j}{N_i}(p_{jk} - p_{ik}), \tag{5.3b}$$

where the average fitness now includes the emigration term, $\overline{r}_i = 1 - M\rho_j + s_i \sum_k (2p_{ik} - 1)$.

5.3.1 Single locus results

In principle, exact expressions for the general solution of the single-locus equilibria can be found. However the results are impenetrable and provide little insight. Instead, we analyse the equilibria by solving the dynamics either exactly for particular cases or, for more general cases, numerically using MATHEMATICA (Wolfram 1999).

Varying habitat frequency ρ

If M is fixed and the habitat frequency varies, two distinct equilibria emerge: either the metapopulation is monomorphic when one habitat is significantly more frequent than the other and its resident allele fixes in all demes, or neither habitat is dominant and the metapopulation is polymorphic. We define a species to be an "evolutionary specialist" when only one allele persists throughout the metapopulation, hence the species succesfully adapts in only one habitat (monomorphic equilibria). If both alleles persist in the metapopulation (polymorphic equilibria), then the species fails to adapt well to either habitat and it is an "evolutionary generalist". We can also define the species as specialists and generalists in an ecological sense (Rosenzweig 1987). If the species has a relatively greater abundance in one habitat (e.g. monomorphic equilibria), we define it to be an "ecological specialist". But if the species is approximately equally divided between both habitats, or neither habitat is dignificantly greater than the other, it is an "ecological generalist" (e.g. polymorphic equilibria).

The monomorphic equilibria represent states in which the species is simultaneously an ecological as well as evolutionary specialist. However, in the polymorphic case, since the population is not approximately equally divided between habitats, the species is only an evolutionary generalist, and not an ecological generalist (see Ronce & Kirkpatrick 2001 for an alternate definition of an ecological generalist species).

The equilibrium population size and allele frequencies of each habitat are functions of ρ and γ and as such are shown as equilibrium solution surfaces in Fig. 5.1. Equilibrium solution curves for fixed γ are also shown.

The equilibrium surfaces change discontinuously along two curves. For the parameter values shown, one curve cuts across the $\gamma - \rho$ plane, whilst the other just clips the corner of the plane for $\rho < 0.5$. These critical curves delineate the parameter values for which the monomorphic and polymorphic equilibria are stable. We obtain a clearer picture of when the metapopulation switches between the two different equilibria by examining the equilibrium curves that result from setting the ratio γ to one, so that the magnitude of selection in both habitats is identical. These curves are the leftmost figures in Fig. 5.1b.

When fraction ρ is below a critical fraction ρ_L, the '$-$' allele, local to habitat 2, invades habitat 1 and fixes: habitat 2 is perfectly adapted and deme sizes within the habitat are well above the intrinsic carrying capacity ($N > 1$). Habitat 1 is maladapted and therefore has a negative growth rate; it is a pseudo sink (Watkinson & Sutherland 1995, Pulliam 2000). The low demic size in habitat 1 is maintained by migration from the source population, habitat 2. As ρ increases, population sizes in both habitats decline. This is due to the combined effect of demography and selection. The average fitness in habitat 2 is depressed whilst the immigration of unfavourable alleles into the habitat increases. Though the average fitness in habitat 1 increases, immigration diminishes due to the shrinking size of demes in habitat 2 and the increased emigration out of habitat 2. Since habitat 1 is a sink and depends on migrants to sustain population numbers, the demic sizes in habitat 1 decrease in parallel. This continues until $\rho = \rho_L$, when the '$+$' allele is able to invade habitat 2. This discontinuous change signals polymorphism in both habitats and a dramatic increase in the demic sizes of habitat 1 with an accompanying decline in demic sizes in habitat 2. The source-sink structure disappears until greater ρ, when the source-sink structure is inverted, see leftmost figures in Fig. 5.1b (Pulliam 1988, Dias 1996). Above a critical frequency ρ_H, the '$-$' allele cannot invade habitat 1 and in a reversal of roles, the '$+$' allele spreads throughout the metapopulation (the alternate

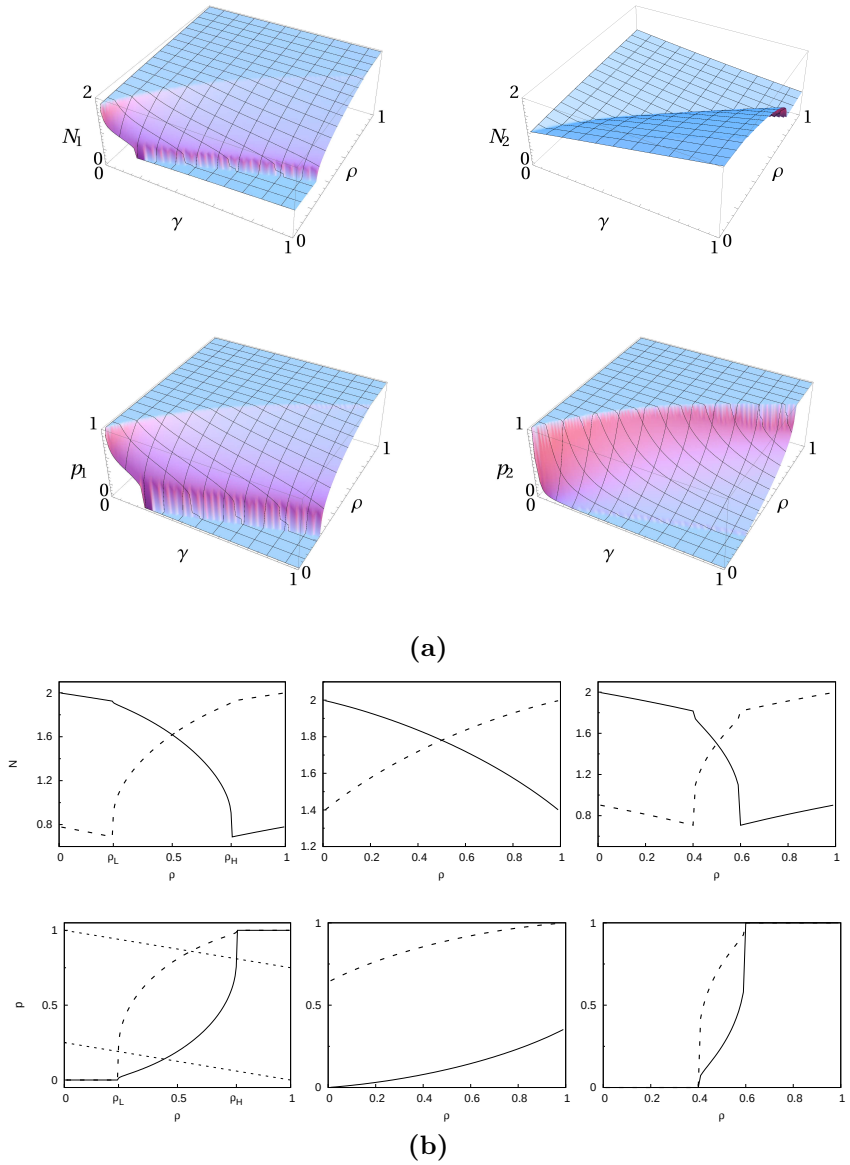


Figure 5.1 *Deterministic equilibria of habitat 1 and 2 for varying ρ and fixed M. (a) Equilibrium surfaces of N (top row) and p (bottom row) for habitat 1 (left) and habitat 2 (right) with $M = 0.5$, $s = 1$ and $\gamma = 1$ (note the ρ axis is different between the two sets of figures). (b) Equilibrium curves of N (top row) and p (bottom row) for habitat 1 (dashed lines) and habitat 2 (solid lines) with $s = 1$, $\gamma = 1$ and migration constants $M = 0.5$ (left column), $M = 0.25$, (middle column) and $M = 0.75$ (right column). The leftmost column figures represent two dimensional slices through the equilibrium surfaces in (a). Source-sink structure emerges when the average fitness falls below zero, $\overline{r}_i < 0$: above the upper dotted line, habitat 2 is a sink and below the lower dotted line habitat 1 is a sink ((a), bottom row, leftmost figure). Critical fractions ρ_L and ρ_H are indicated.*

specialist state).

Source-sink structure in the metapopulation exists provided one of the two habitats has zero or negative average fitness. The criterion for this is $2s_ip_i \leq m\rho_i + s_i - 1$ and is indicated in the leftmost, bottom figure in Fig. 5.1b.

Irrespective of the fractional size of either habitat, conditions for polymorphism are favourable at all fractional sizes, except when $\rho \sim 0$ and $\rho \sim 1$ see Fig. 5.1b (middle figures). In contrast, under strong migration, though both populations are well-mixed, one allele fixes in the metapopulation. This is because the larger habitat has a greater bias, no matter how slight, in the migrant pool: adaptation is facilitated by an ecological factor, the habitat size.

As expected, for the special case $\gamma = 1$, the equilibria are symmetrical about $\rho = 1/2$ in the sense that $N_1(\rho) = N_1(1-\rho)$ and $p_1(\rho) = p_2(1-\rho)$. The symmetry is broken when $\gamma \neq 1$, as we shall see when we examine the critical fractions ρ_H and ρ_L below.

Polymorphic equilibrium

For habitat frequencies between the critical values, $\rho_L < \rho < \rho_H$, exact solutions of the equilibria cannot be found. The implicit solutions are,

$$\begin{aligned} N_i &= \frac{\overline{r}_i + \sqrt{\overline{r}_i^2 + 4M\rho_j N_j}}{2}, \\ p_i &= \frac{(N_i s_i - M\rho_j N_j) \pm \sqrt{(N_i s_i - M\rho_j N_j)^2 + 4NMs_i\rho_j N_j}}{2N_i s_i}. \end{aligned} \tag{5.4}$$

Only when the metapopulation is equally divided between both habitats and selection is equally strong, $\gamma = 1$, is there no bias in the migrant pool towards either allele.

Monomorphic equilibria

The monomorphic equilibria are given by

$$N_i = \frac{1}{M\rho_j} N_j \Big(N_j - \overline{r}_j \Big), \tag{5.5a}$$

$$\begin{aligned} p_i &= 0,\ 1 \\ &= p_j. \end{aligned} \tag{5.5b}$$

If ρ is at least as large as ρ_L, the '+' allele can invade habitat 2. Similarly, if ρ is at most as great as ρ_H, the '−' allele can invade habitat 1.

The critical habitat frequencies can be found from the Jacobian $\mathbf{J}$ of the dynamics, $J_{lk}\partial_{y_k}\partial_{y_l}\dot{y}_k$ (Strogatz 2004). At these fractional sizes, the equilibrium

curves are discontinuous and at least one of the eigenvalues of **J** is zero. Since the determinant of a matrix is a product of eigenvalues, at the critical fractions, the determinant of **J** is zero. This gives the following equation for the critical habitat sizes, $\rho_c \in \{\rho_L, \rho_H\}$,

$$\rho_c = \frac{\gamma M N_1^2 + (-1)^{1+p} s\gamma N_1 N_2}{M(\gamma N_1^2 + N_2^2)}, \tag{5.6}$$

where $p = 0, 1$ for $\rho_c = \rho_L$ and $\rho_c = \rho_H$ habitat fractions respectively. By substituting the expression for N_1 and N_2 from Eq. 5.7a, both of which depend on ρ_c, the above expression can be numerically solved to give the critical habitat frequencies.

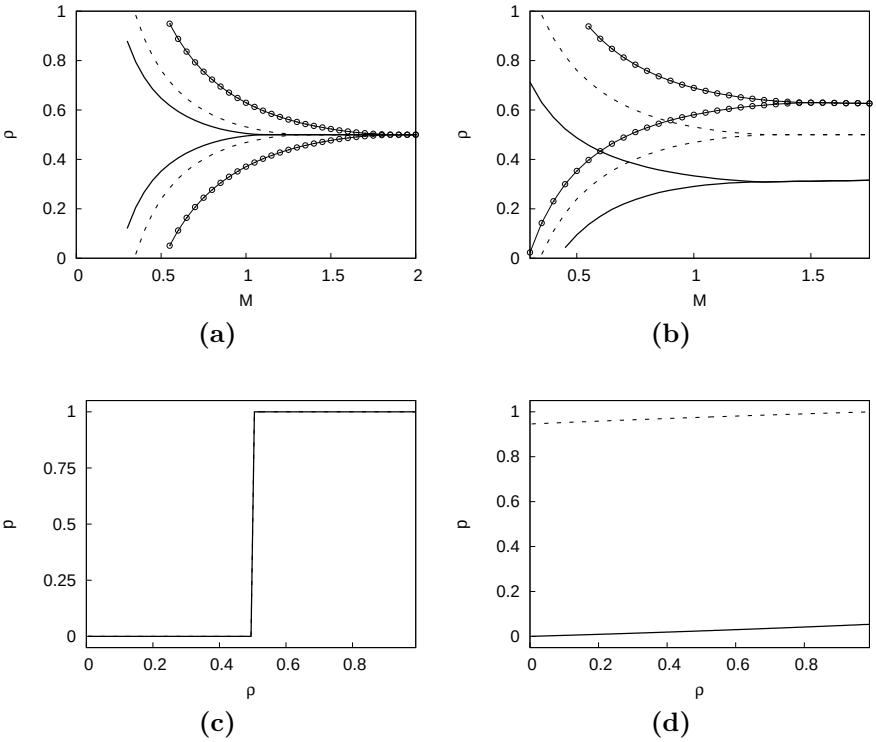


Figure 5.2 *Critical habitat frequencies ρ_H, ρ_L. Top row, ρ_H (upper curves) and ρ_L (lower curves) for different values of M. (a) $s = 0.5$ (solid lines), $s = 0.75$ (dashed lines) and $s = 1$ (with circles). (b) $\gamma = 0.5$ (solid), $\gamma = 1$ (dashed) and $\gamma = 1.5$ (with circles). (c & d) The equilibrium curve for p against ρ is also shown for habitat 1 (dashed lines) and habitat 2 (solid lines). Either (c) the critical fractions collapse to one value for high migration ($s = 0.01$, $M = 10$), or (d) they disappear entirely for strong selection ($s = 1$, $M = 0.05$). Unless otherwise stated, $\gamma = 1$, $s = 1$.*

Using (5.4) in (5.6), critical frequencies ρ_L and ρ_H can be calculated numerically. The results are shown in Fig. 5.2. Above the lower curves (ρ_L), below the upper curves (ρ_H) and for low migration rates where neither curve

exists, the metapopulation is polymorphic.

If $\gamma = 1$, the critical fractions satisfy $\rho_L = 1 - \rho_H$ and are symmetrical about $\rho = 1/2$, see Fig 5.2a. As selection increases, for a given migration constant M, the critical fractions diverge and the range of habitat frequencies for which the metapopulation is polymorphic increases.

If one allele is (relatively) weakly selected within its favourable habitat so that $\gamma \neq 1$, one of the two critical fractions may disappear and only one of the two alleles can invade, see Fig. 5.2b. Polymorphism is guaranteed at any habitat size if both critical frequencies disappear (and $0 < \rho < 1$). Notice that the curves are no longer symmetric about $\rho = 1/2$.

For large migration rates M, the difference between ρ_H and ρ_L decreases until both fractions collapse to one value. In this case, polymorphism is no longer possible as both alleles successfully invade the respective foreign habitat and the species switches between specialist states, see Fig. 5.2c. For strong selection (or increased isolation) both alleles persist in the metapopulation and each habitat is well adapted, suffering only small migration loads, see Fig. 5.2d.

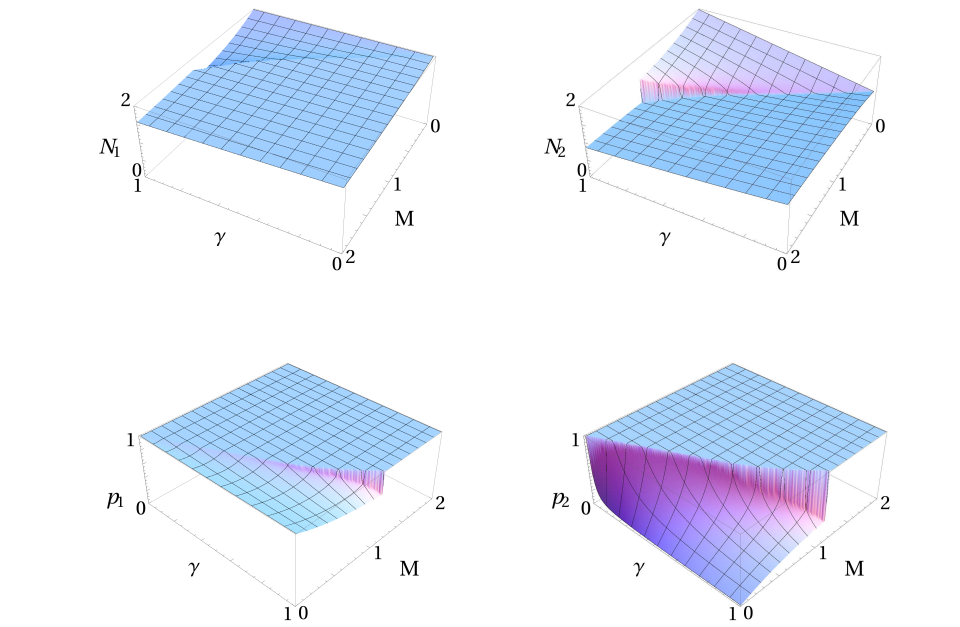
We have considered equilibria for the case where the migration constant is fixed and habitat frequency ρ and ratio γ vary. Next, we keep ρ fixed and examine equilibria for vary migration constant M.

Varying migration constant M

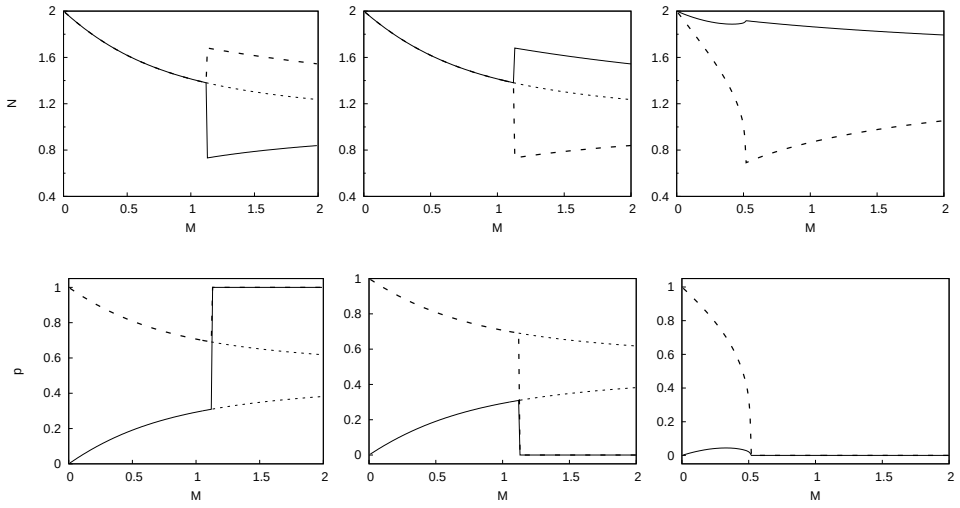
For fixed habitat sizes, the metapopulation is polymorphic for migration constants below critical value M^*, and monomorphic otherwise. This is the familiar threshold between adaptation and the 'swamping' effect of migration (see Ch. 2). We can further divide the equilibria according to whether the habitats are of equal size ($\rho = 1/2$) or not.

The equilibrium population size and allele frequency for fixed ρ are shown in Fig. 5.3. A critical curve marks a discontinuous change in the equilibrium surfaces. This critical curve corresponds to the critical rate M^* as a function of γ, see Fig. 5.3a.

For the special case of $\rho = 1/2$ and $\gamma = 1$, the habitats are symmetrical for migration rates below M^*, see leftmost figures in Fig. 5.3b. By symmetrical we mean that the population sizes in each habitat are equal and $p_i = 1 - p_j$: the population adapts successfully to both habitats and is equally abundant in both habitats (though since there is some migration between the two habitats, the population suffers a migration load in each habitat). The species is a generalist,



(a)



(b)

Figure 5.3 *Deterministic equilibria of habitat 1 and 2 for varying M and fixed ρ.(a) Equilibrium surfaces of N (top row) and p (bottom row) for habitat 1 (left) and habitat 2 (right) with $\rho = 0.5$ $s = 1$, $\gamma = 1$ and initial conditions $N_0 = 1$, $p_0 = 0.9$ (the M-axis is different between the two sets of figures). (b) Equilibrium curves of N (top row) and p (bottom row). Left and middle columns: stable equilibria of habitat 1 (dashed lines) and habitat 2 (solid lines) are shown for $\rho = 0.5$, and initial conditions $N_0 = 1$, $p_0 = 0.9$ (left column), $p_0 = 0.1$ (middle column). Dotted curves indicate unstable equilibria. Right column, $\rho = 0.25$. Unless otherwise stated, $s = 1$ and $\gamma = 1$.*

in both the ecological and evolutionary sense of the term. The equilibrium population sizes and allele frequencies are

$$N_i = N_j = 1 + s_i(2p_i - 1), \tag{5.7a}$$

$$p_i = \frac{(s_i - M) \pm \sqrt{(s_i - M)^2 + 2Ms_i}}{2s_i} \tag{5.7b}$$

$$= 1 - p_j.$$

For migration constants above M^*, a specialist equilibrium emerges. In the specialist state, similar to the monomorphic equilibria described in the previous section, one allele fixes throughout the metapopulation. This habitat is the more abundant habitat and once again a source-sink structure emerges: the invaded habitat is maladapted and the demic size is correspondingly small. Furthermore, the specialist equilibrium is bistable: the initial state of the metapopulation dictates which allele fixes, see Fig. 5.3b. For the special case $\rho = 1/2$, $\gamma = 1$, $M^* = \frac{1}{3}\left(2 + \sqrt{1 + 3s^2}\right)$. Note this is proportional to s, as expected from soft selection models (see §2.2).

When the habitats are unequal in fractional size, $\rho \neq 1$, the generalist/specialist description is no longer appropriate, since the habitats are not equally abundant. Furthermore, only at very low migration rates can the species evolve to be generalist: even for migration rates below $M^*\rho$, one habitat rapidly becomes maladapted as M increases. This habitat is the less frequent habitat. In other words, (metapopulation) demography constrains evolution. Which allele fixes at high migration rates simply depends on which habitat is more frequent or which allele has the stronger selective advantage i.e. the '+' allele fixes if $\rho > \frac{1}{2}$ or $\gamma < 1$, and conversely for the alternate allele, see Fig. 5.3b. The demic sizes are still symmetric about $\rho = 1/2$, in the sense that the demic size in habitat 1 for ρ_a is the same as the demic size for habitat 2 if $\rho_b = 1 - \rho_a$ i.e. $N_1(\rho_a) = N_2(\rho_b)$.

Irrespective of the value of ρ, as the migration constant increases above rate M^*, the demic sizes within both habitats appear to converge to the same value. This is because the rate of emigration increases with M; the source habitat incurs a demographic cost, even though is 'perfectly' adapted.

We have examined the different equilibria that arise when either the migration constant M or habitat fraction ρ is varied. We have found that genetic polymorphism is favoured when either M is small, s is large, one allele has a greater absolute advantage i.e. $\gamma \neq 1$, and habitat sizes are intermediate. Local adaptation in both habitats is possible if habitats are almost completely isolated ($M \ll 1$) or there is no bias toward either allele in the metapopulation and $\gamma = 1$. Then the species evolves to be an all-round generalist species.

We next consider the effect of a catastrophe on the metapopulation, before investigating multilocus evolution.

The effect of a catastrophe

What happens if a habitat is subject to a catastrophe, such as a reduction in size? To investigate the effect of a catastrophe, we follow the procedure outlined in Ronce & Kirkpatrick (2001). The habitats evolve for 100 generations, with habitat 1 initially as a source population, and habitat 2 as the sink population. After 100 generations, habitat 1 suffers a reduction in size i.e. ρ decreases.

The results are shown in Fig. 5.4. We decreased the fractional size of habitat 1 from $\rho = 0.9$ to two different sizes, $\rho = 0.4$ and $\rho = 0.1$. The consequent reduction of immigration into habitat 2 enables the habitat to evolve. If habitat 1 suffers a mild catastrophe, that is, a small decrease in size, adaptation occurs in both habitats to some degree and the source-sink structure disappears, see Fig. 5.4a. This is an example of niche evolution (Holt 2009). On the other hand, if the catastrophe is severe enough and ρ is drastically reduced, the '$-$' allele successfully invades the metapopulation. Furthermore, habitat 1 becomes a sink and the initial source-sink structure is inverted, see Fig. 5.4b.

Notice that immediately after the catastrophe the population size in habitat 2 is temporarily depressed even further, before rebounding to higher values ($N > 1$). During this transitional period, habitat 2 is more vulnerable after the catastrophe than before the catastrophe occurred.

We come to the end of the single locus results and to conclude the deterministic section, next examine the effect that multiple loci have on the evolution of the two habitats. The multilocus case is a simple extension of the single locus model, therefore we expect the results to be qualitatively similar. However, the conditions for genetic polymorphism critically depend on locus numbers.

5.4 Multiple loci

We can easily generalise the single locus results to multiple loci. We assume loci are symmetric and independent and so the average fitness of habitat i is now $\overline{r}_i = 1 - M\rho_j + ns_i(2p_i - 1)$, where n is the number of loci contributing to the trait (j refers to the other habitat). Aside from this small difference, the dynamics are as given in 5.3.

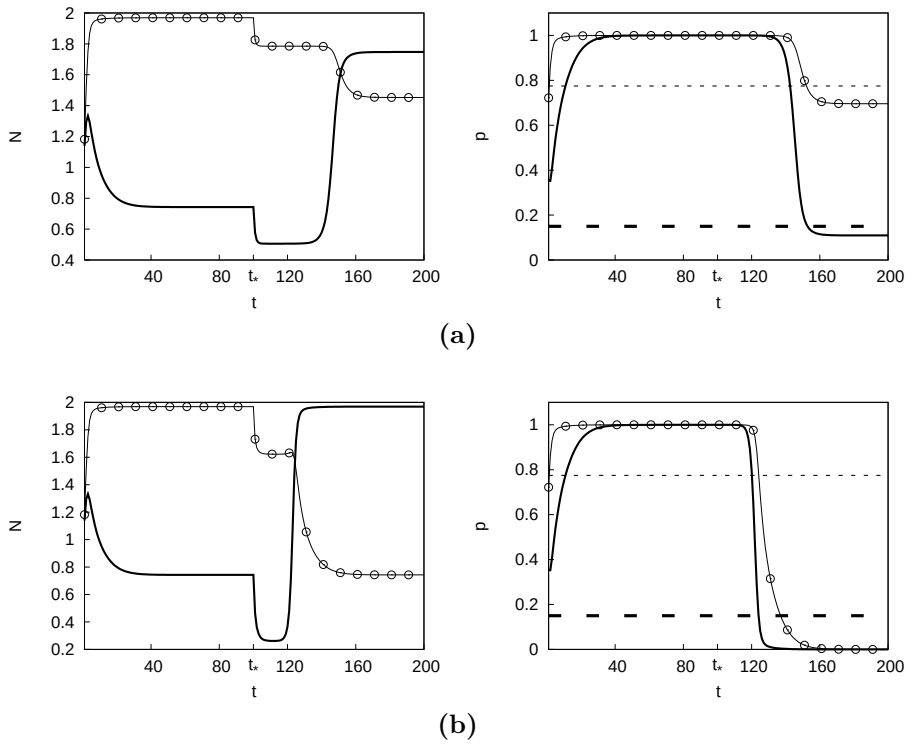
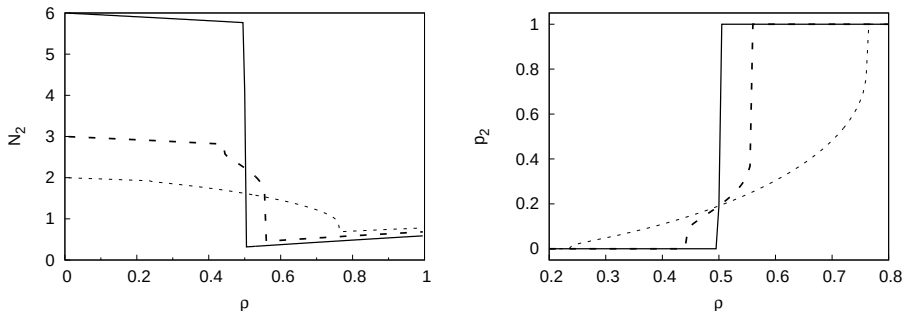


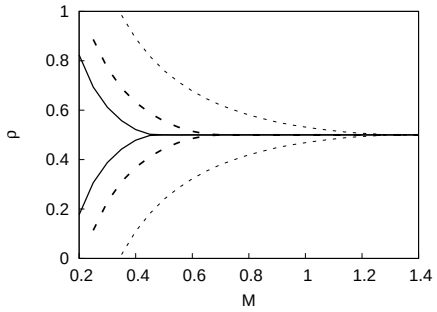
Figure 5.4 *The effect of a catastrophe (reduction in ρ) on demic size N (left) and allele frequency p (right) when initially $\rho = 0.9$, $N_0 = 1$ and $p_0 = 0.5$ with $s = 1$, $\gamma = 1$ and $M = 0.5$. The fractional size of habitat 1 (with circles) is reduced to (a) $\rho = 0.4$, and (b) $\rho = 0.1$. Initially, habitat 2 is a sink. Habitat 1 is a sink if p_1 is below the thick dashed line. Habitat 2 (solid lines) is a sink if p_2 is above the light dashed line.*

As discussed in Ch. 3, under the assumption of symmetry, the number of loci n simply features as a multiplicative factor in the average fitness. This means that selection per locus, s_i, can be weaker the greater the number of loci are involved whilst the effective selection coefficient, ns_i, remains fixed.

The equilibrium multilocus results are shown in Fig. 5.5 for varying ρ. As expected, the equilibrium population size and allele frequency curves for fixed migration rate M are qualitatively similar to the single locus results: there are two monomorphic regimes for small and large ρ and a polymorphic phase for intermediate ρ, see Fig. 5.5a. However, as locus numbers increase, the range of ρ for which polymorphism can exist in the metapopulation decreases, and the transition between monomorphic states becomes sharp. This means that the source-sink structure is inverted over a smaller migration range as n increases. This can occur for modest loci numbers, for example $n > 5$. The sharp transition is similar to the sharp collapse of the high state (i.e. collapse of bistability) at high migration rates in the deterministic multilocus model in Ch. 3, cf. 3.4.



(a)



(b)

Figure 5.5 *Deterministic multilocus results for locus numbers $n = 1$ (dotted lines), $n = 2$ (dashed lines) and $n = 5$ (solid lines).(a) equilibrium population size (left) and allele frequency (right).(b) Critical fractions ρ_L (lower curves) and ρ_H (upper curves). Unless otherwise stated, $s = 1$, $\gamma = 1$ and $M = 0.5$.*

The critical fractions ρ_L and ρ_H are shown in Fig. 5.5b. It is clear from the figures that as locus numbers increase, the critical fractions approach $\rho = 1/2$, and the conditions for polymorphism become restricted. Large locus numbers have a similar effect on the metapopulation as high migration (cf. Fig. 5.2c).

When a large number of symmetric loci contribute to fitness, the maximum and minimum average fitnesses become polarised. Therefore a habitat is either close to perfect adaptation or close to total maladaptation. In contrast to the single locus case, habitat 2 is in a sense buffered from the degrading effect of maladaptive gene flow, whereas habitat 1 becomes highly maladapted. This is true so long as ρ is small enough that the bias in the migrant pool is towards habitat 2 (i.e. for $\gamma = 1$, $\rho < \frac{1}{2}$). Once ρ is large enough (habitat 2 is rare) habitat 2 is overwhelmed by maladaptive immigrants and the metapopulation switches between monomorphic states. Therefore the evolutionary effect of multiple loci is overcome by demography and migration (since $M\rho$ is the true migration rate).

We have examined the deterministic dynamics of the island model and considered the effect of a single locus and multiple (symmetric) loci. We have

found that above or below a critical fraction (for fixed migration rates) or a critical migration rate (for fixed habitat frequencies), one allele is guaranteed to invade the rest of the metapopulation, provided migration is intermediate and assuming selection is relatively weak. Whether polymorphism exists in the metapopulation depends crucially the number of loci contributing to trait and the relative strength of migration and selection (per locus).

We now investigate the effect of genetic and demographic drift on the metapopulations.

5.4.1 Stochastic demes

We examine how stochastic fluctuations alter the deterministic results of the previous section. We focus on a single locus; extending the formalism below to multiple loci is straight-forward. Moreover, the qualitative results of the single locus case will carry over the the symmetric multilocus case.

When demes follow stochastic dynamics, the population size and allele frequency of habitat i are distributed about mean values $\langle N_i \rangle$ and $\langle p_i \rangle$ respectively. Averages taken across a group of demes, denoted $\langle x \rangle$, vary discontinuously, making large jumps from one generation to the next, if the number of demes is small. On the other hand, in the limit of an infinite number of demes these averages vary smoothly and a diffusion description is applicable. We therefore make the important assumption that the metapopulation consists of an infinite number of demes (Wright 1931, Barton & Rouhani 1993*a*). We also assume that changes within a deme are much faster than changes on the scale of the whole metapopulation. At a given instant, habitat i can then be described by a distribution of states $\Psi_i(\mathbf{x}_i|\hat{\mathbf{y}})$, conditional on the instantaneous state of the migrant pool (determined by the set of migrant averages $\hat{\mathbf{y}}$). Since all demes contribute to the migrant pool, as explained in §5.2, $\hat{y} = \langle y \rangle$, where the latter is an average across the entire metapopulation. Again, we consider only two habitats. Though individual demes fluctuate, distribution $\Psi_i(\mathbf{x}, t|\hat{\mathbf{y}}, t')$ evolves deterministically until the migrant/metapopulation averages converge to fixed equilibrium values, $\hat{y} = \langle y \rangle = \langle y \rangle^*$. Barton & Rouhani (1993*a*) have proved that this convergence occurs for an infinite number of demes under soft selection.

The stochastic dynamics of a deme within habitat i are

$$dN_i = \Big(N_i(\overline{r}_i - N_i) + M\langle N\rangle\Big)dt + \sqrt{\frac{N_i}{\zeta}}dW_{N_i}, \tag{5.8a}$$

$$dp_i = \Big(s_i p_i(1-p_i) + \frac{M(\langle Np\rangle - \langle N\rangle p_i)}{N_i}\Big)dt + \sqrt{\frac{p_i(1-p_i)}{2\zeta N_i}}dW_{p_i}, \tag{5.8b}$$

At any given instant, to habitat i the rest of the metapopulation will appear fixed and the instantaneous migrant averages can be treated as constant. The stochastic dynamics are such that they can be written in potential form. Using the Fokker-Planck formulation, the instantaneous distribution of states across habitat i, $\Psi_i(\mathbf{x}|\hat{\mathbf{y}})$ given migrant averages $\hat{y}$ is

$$\Psi_i(\mathbf{x}_i \,|\, \hat{\mathbf{y}}) = \frac{1}{Z_{0,i}}\frac{e^{2\zeta U_i}}{N_i p_i q_i} \tag{5.9}$$

$$= \frac{1}{Z_{0,i}} N_i^{2\zeta M\langle N\rangle - 1} e^{2\zeta N_i \overline{r}_i - \zeta N^2} p_i^{4\zeta M\langle Np\rangle - 1}(1-p_i)^{4\zeta M\Big(\langle N\rangle - \langle Np\rangle\Big) - 1}, \tag{5.10}$$

where

$$U_i = \overline{r}_i N_i - \frac{1}{2}N_i^2 + M\langle N\rangle \ln N_i + 2M\eta\Big(\langle Np\rangle \ln p_i + \Big(\langle N\rangle - \langle Np\rangle\Big)\ln q_i\Big). \tag{5.11}$$

and $Z_{0,i}$ is the normalisation. This is similar to the single locus stationary distribution of the stochastic mainland-island model from the last chapter, see §4.3; the stationary probability distribution of one island evolving under the same stochastic dynamics is equivalent to the distribution of states across a large ensemble of demes. In the limit of weak drift, high selection $\zeta s \gg 1$ and high migration $\zeta M\rho \gg 1$, $\langle Np\rangle = \langle N\rangle\langle p\rangle$, the covariance between N and p across all demes tends to zero and the stochastic dynamics will closely mirror the deterministic dynamics to leading order in ζ^{-1}

To determine the stationary distribution, $\Psi_i(\mathbf{x} \,|\, \{\langle \mathbf{y}\rangle^*\})$, we need to find $\langle \mathbf{y}\rangle^*$ by solving

$$\langle N\rangle = \sum_{i=1}^{2} \rho_i \Big(\int N_i \Psi_i(\mathbf{x}_i \,|\, \hat{N}p,\, \widehat{Np})\, \mathrm{d}\mathbf{x}_i \Big) \tag{5.12a}$$

$$\langle Np\rangle = \sum_{i=1}^{2} \rho_i \Big(\int N_i p_i \Psi_i(\mathbf{x}_i \,|\, \hat{N},\, \widehat{Np})\, \mathrm{d}\mathbf{x}_i \Big). \tag{5.12b}$$

The stationary averages are found by the following numerical procedure. First, averages $\langle y\rangle$ are numerically calculated for each habitat. Second, these averages are then substituted back into the distribution of states of each habitat, $\Psi_i(\mathbf{x}_i|\hat{\mathbf{y}})$ using $\hat{y} = \langle y\rangle$. These steps are repeated until the migrant/metapopulation averages converge to equilibrium values.

If either habitat is rare, $\rho \sim 0, 1$, the metapopulation is similar to the stochastic mainland-island model from the last chapter (emigration from the smaller habitat would have a negligible effect on the rest of the metapopulation). We therefore focus on the case where both habitats are equally frequent and selection is opposite but equal ($\rho = 1/2$, $\gamma = 1$).

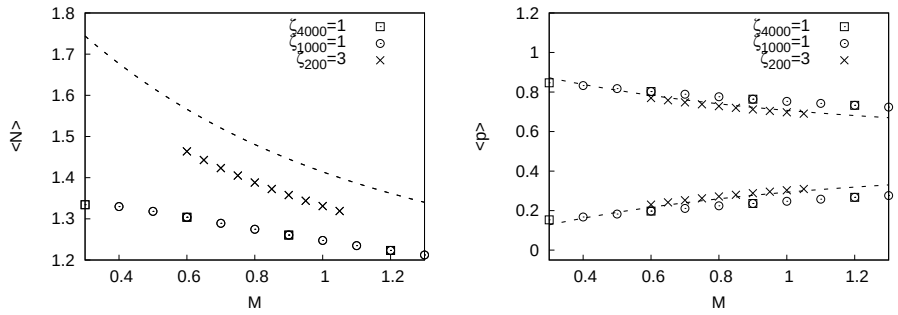
Since the metapopulation evolves deterministically, in principle we should be able to determine the equilibrium fixed points of (5.12). For some function $F(x)$, the fixed point x' satisfies $\Big(\partial F(x)/\partial x\Big)|_{x=x'} = 1$. Noting that $\rho = 1/2$, if we therefore take the derivative of $\langle N\rangle$ (5.12a) with respect to $\hat{y}$, where $\hat{y} \in \{\hat{N},\, \widehat{Np}\}$, we obtain the general condition (cf. Eq. 6 and 16a in Barton & Rouhani (1993a))

$$\sum_{i=1}^{2} \mathrm{Cov}\Big(N_i,\, h_i\Big) = \frac{1}{\zeta M}. \tag{5.13}$$

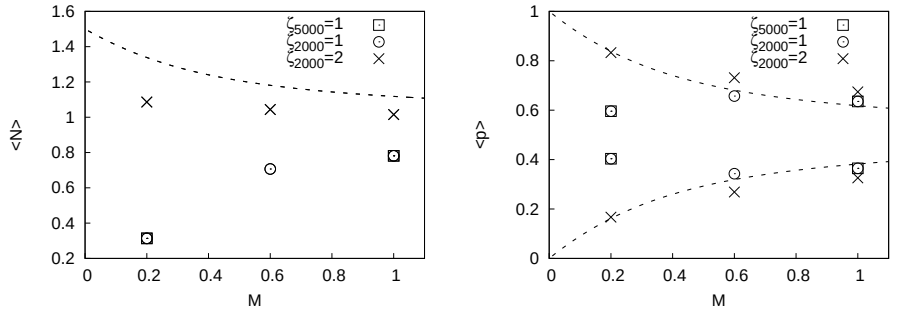
where $h_i = \frac{1}{M}\partial_{\hat{y}} U_i$ and $\mathrm{Cov}(a,\, b)$ is the covariance of a and b. There is no simple interpretation of this condition or the potential derivative, h_i. However, it does suggest that the covariance of N_i and h_i within each habitat tends to zero the weaker drift is or the stronger migration is (ζM increases), as asserted earlier.

Numerical results

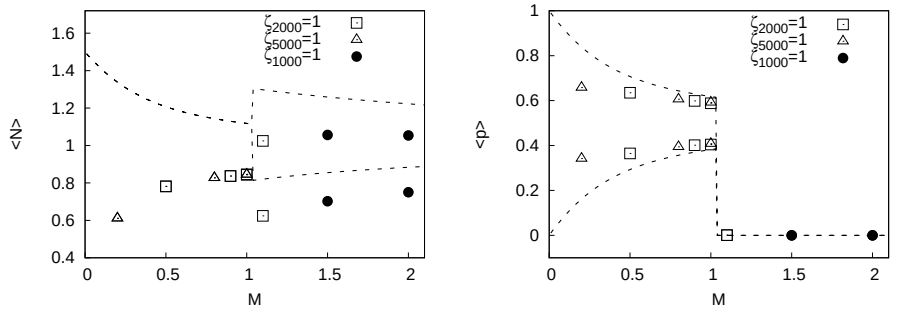
Numerical results are shown in Fig. 5.6 for $\rho = 1/2$ and moderate ($\zeta s,\, 2\zeta M < 1$) to weak ($\zeta s,\, 2\zeta M \geq 1$) drift. The figures show the stationary average demic size and average allele frequency in each habitat. From the figures we see that, as expected, for high migration and weak drift the average demic sizes and allele frequencies are qualitatively similar to the deterministic results in §5.3.1, cf. Fig. 5.3, see e.g. $\zeta = 2$ curves in Fig. 5.6a. (This is not true for the average



(a) $s = 1$, *i2*



(b) $s = 0.5$, *i2*



(c) $s = 0.5$, *i3*

Figure 5.6 *Equilibria for stochastic demes for fixed ρ: equilibrium average demic size (left) and allele frequency (right) in each habitat for the symmetric case with $\rho = \frac{1}{2}$, $\gamma = 1$ for $s = 0.5$, 1, various ζ and number of iterations (shown as subscripts on ζ in figures). Where there are two sets of the same symbols, the upper set is for habitat 1, otherwise the results are identical for both habitats. The deterministic equilibria were used as initial conditions. The initial conditions for the deterministic equilibria are identical for both habitats. Set i2 is $N_0 = 1$, $p_0 = 0.5$, set i3 is $N_0 = 1$, $p = 0.1$. Deterministic equilibria are also shown (dashed lines), cf. Fig. 5.3. (In* MATHEMATICA*, the precision goal was set to 5, the working precision 15, max recursion 20. At best, the error in data y will be approximately $|y| \times 10^{-5}$ (Wolfram 1999).)*

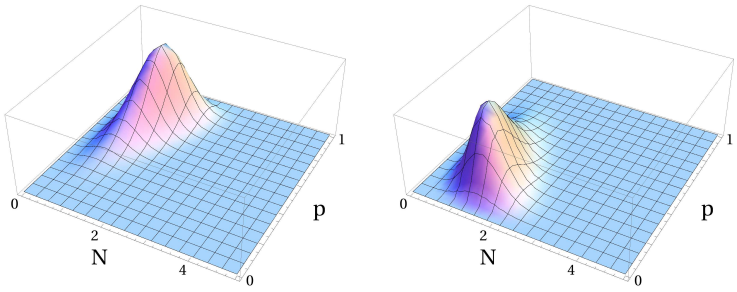
allele frequencies, when $\zeta = 2$ in Fig. 5.6a. We suspect that this is because convergence to equilibrium may not have occurred.) We expect the stochastic results to converge to the deterministic results to leading order in ζ^{-1}.

However, the demic sizes are much lower than the deterministic predictions, especially for weak selection. This is due to demographic stochasticity; a certain fraction of demes in each habitat will either be extinct or have low demic sizes. This is not an artifact of convergence, since we have checked certain data points for significantly longer iteration runs and found the result to be indistinguishable, e.g. see $\zeta_{1000} = 1$ and $\zeta_{4000} = 1$ in Fig. 5.6a, where the subscripts indicate the number of iterations.

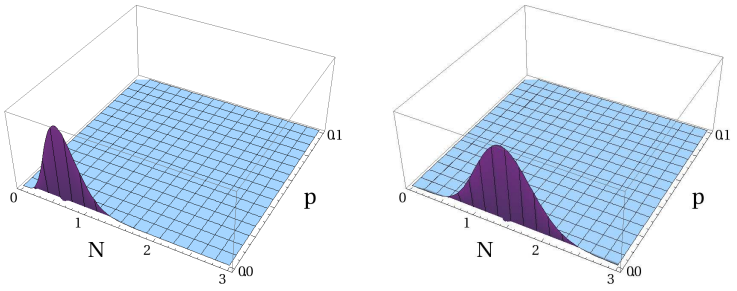
For low migration and selection, when drift dominates ($\zeta s \ll 1$, $\zeta M \ll 1$), just over half the demes will fix for the favoured allele in each habitat and just under half will not. The average allele frequency in each habitat tends to 1/2 and the average demic size is correspondingly low (due to high migration loads) as selection and migration decrease, see Fig. 5.6b. This could lead to extinction, even though maladaptive gene flow between habitats is reduced. Ironically, as demes become increasingly isolated, on average the species is increasingly more generalist, since both alleles tend to be equally present in each habitat. This is not because of a well-mixed metapopulation, but because habitat sizes are equal and the strength of selection does not vary between habitats. If habitat sizes were unequal, at low migration rates the more favoured allele of the more frequent habitat would be fixed in more demes. Similarly for the (absolute) selective advantage of an allele.

The results suggest that the critical migration rate above which one allele fixes throughout the metapopulation is very similar to the deterministic value, see Fig. 5.6b. This is somewhat surprising, since we would expect the discrepancy between the stochastic and deterministic critical rate to be at least to leading order in ζ^{-1}. In fact, we would expect the critical rate to be lower in the stochastic case, since extinction and small population sizes should make the conditions for polymorphism less favourable than the deterministic case, as has been found for soft selection models (Barton & Whitlock 1997). A more thorough numerical treatment of the stochastic model is required.

Figure 5.7 shows the habitat distributions, Ψ_1 and Ψ_2, see (5.9). At low migration and weak selection, both distributions have the characteristic U-shape of drift-dominated distributions. As migration increases, the distributions peaks approximately about the corresponding deterministic equilibria, see Fig. 5.7a. As migration increases further, one allele spreads throughout the metapopulation, depending on the initial state of the metapopulation. In Fig. 5.7b, this is the '$-$' allele. Notice that Ψ_2 is centered about a larger demic size than Ψ_1, since demes within habitat 2 are fitter.



(a) $M = 1$



(b) $M = 1.5$

Figure 5.7 *Distributions* Ψ_1 *(left) and* Ψ_2 *(right) for* $s = 0.5$, $\zeta = 2$, *using deterministic values for* $\langle N \rangle$, $\langle Np \rangle$.

5.5 Summary and discussion

We have examined an infinite island model and determined the conditions favourable for local adaptation or genetic polymorphism under both deterministic and stochastic dynamics, for one locus and many loci. The basic result from the classic multi-niche models (see Ch. 2) hold: for one locus, if the habitat frequency (or quality, see Dias (1996)) can vary, then local adaptation and specialisation is likely the more frequent the habitat, $\rho_i \gg 1/2$, or the more strongly selected the local allele, $|s_i| \geq 1$. For fixed ρ, isolation promotes genetic polymorphism and hence adaptation to local conditions. This is a classic result, which is unaltered by the inclusion of population dynamics. We find that the distinction between "ecological" and "evolutionary" generalists (specialists) becomes important.

Generalists and specialists: the only case for which the species' is an eco-evolutionary specialist or generalist is in the highly symmetric case of equal habitat sizes when selection only varies in direction between the two habitats. This case is similar to Ronce & Kirkpatrick (2001) model, in which trait

evolution under stabilising selection was studied. They find that both generalist and specialist equilibria are stable at very low migration rates, otherwise the population is a specialist at intermediate migration rates. However, at very high migration rates, the population becomes a generalist. In contrast, we find that at low migration rates the species is a generalist and is able to adapt to some degree. At greater migration rates, specialism evolves and source-sink structure emerges. Isolation is favourable to genetic polymorphism, as has been shown by single-locus soft selection models (see Ch. 2). Which allele spreads in the metapopulation is a matter of history and depends on the initial state of the metapopulation before migration between habitats begins. We have found that the 'hard demographic' cost of two-way migration is such that even if a habitat is optimally adapted, its demic size can be reduced due to high emigration. Crucially however, the interaction between selection and gene flow gives rise to a critical migration threshold, M^*, below which adaptation in both habitats is possible. The inclusion of population dynamics does not alter this basic feature of classical population genetic models.

Habitat size: if habitat frequencies can vary, genetic polymorphism is possible between critical fractional sizes ρ_L and ρ_H. If selection varies only in direction, these critical fractions are symmetric about $\rho = 1/2$, similar to the multi-niche models discussed in Ch. 2. However, if the strength of selection varies between habitats, one critical fraction may be lost and the conditions for polymorphism increase (Bulmer 1972).

Interestingly, Van Tienderen (1991) finds that in a quantitative genetic model of trait evolution, as the frequency of one habitat is gradually reduced, say due to environmental change, the population shifts between specialising on that habitat, being a generalist, to specialising on the remaining more frequent habitat. Van Tienderen notes that this is similar to periods of stasis between which there is a transitional period of rapid change and high phenotypic variation. This is strikingly similar to what happens in the deterministic island model as ρ is increased: the species moves between monomorphic states via a generalist phase.

Catastrophes: a catastrophe may not only change the size of a habitat but it may affect the quality of the habitat in other ways. This may be reflected by a sudden change in habitat frequency or by the relative selective advantage of each allele. We have shown that source-sink inversion can occur should the metapopulation suffer a severe catastrophe. A decrease in γ will have a qualitatively similar effect as the reduction in ρ, because both affect the average fitness of each habitat. Similarly, Ronce & Kirkpatrick (2001) have

considered the effect of habitat disturbances on the metapopulation and found that extinction can occur. This has important consequences for conservation management, especially in the light of habitat destruction due to anthropogenic disturbances (Hanski 1999).

Multilocus evolution: similar to the mainland-island model in Ch. 3, a large number of loci constrain polymorphism, polarising the population sizes. This happens for low locus numbers. Local adaptation is favoured by multilocus evolution in the sense of habitat specialisation, but at the cost of losing genetic variation, which may impair the species' ability to adapt in the future.

Stochastic effects: demographic and genetic drift do not alter the qualitative results of the deterministic model if drift is weak and selection and migration are relatively moderate to strong. However, the expected demic sizes in each habitat are greatly reduced by drift. What is surprising however, is that drift does not appear to alter the critical migration rate above which adaptation to both habitats is no longer possible. Barton & Whitlock (1997) however found that genetic drift and frequent extinctions greatly reduces the conditions for genetic polymorphism. For low migration and selection, the demic sizes in both habitats are much less than the intrinsic carrying capacity. Furthermore, both habitats become increasingly maladapted as migration decreases. This is because drift dominates the dynamics at such low migration values and an approximately equal number of demes are divided between fixation of either allele.

Barton & Rouhani (1993*a*) considered polygenic trait evolution under soft, disruptive selection in an island model (see also Barton & Rouhani (1993*b*)). Our stochastic results are in qualitative agreement with their results, in that above a critical migration rate one adaptive state dominates the metapopulation. This is not surprising: disruptive selection across a homogeneous environment is similar to directional selection across two environments, except it does not have the associated cost of habitat specialisation. Demographic and genetic drift can reduce the average population size of local demes, although their genetic evolution may remain relatively unperturbed surprisingly. Once drift dominates migration and selection, the habitats become increasingly homogeneous through repeated extinction, infrequent colonizations and the depletion of genetic variation. Although the species' appears to become increasingly generalist, in the evolutionary sense, the average demes are severely reduced. It may be that the entire island population could go extinct under demographic fluctuations without frequent recolonization.

Issues and future work: the island model suffers from the same unrealistic assumptions as the deterministic and stochastic mainland-island models and therefore we do not discuss them here. In addition to these idealised assumptions, migration between demes is treated as random, which is unrealistic. In nature, migration is spatially limited. Therefore an obvious extension of the model would be to investigate the effect of limited migration on local adaptation. For example, instead of an infinite number of demes, a finite number of demes with adjacent migration between demes is one possibility (stepping stone model). Abandoning the assumption of a large number of demes will complicate the dynamics however, since this assumption enabled us to utilise the stationary distribution of states from the one-island model of the last chapter.

Barton & Rouhani (1993*b*) investigated the effect of hard selection in the third phase of Wright's shifting balance theory (Wright 1931, 1932, Turelli et al. 2001; see also Ch. 2). However, they only implicitly accounted for changes in population size. One possible avenue for further work would be to extend their model by including explicit population dynamics (as suggested in Barton & Rouhani 1993*b*). This would lead to considerations about speciation, which is beyond the scope of this thesis.

There are other possible extensions. For example, what would happen if the density dependence in each habitat were different or if ecological parameters (other than ρ) varied between habitats? What would happen if the environment were to affect evolution and vice versa? How important would linkage disequilibria be? What if selection varied in time also? Clearly the scope for future work is immense.

Finally, in the next chapter I summarise the results of the last three chapters and discuss the overall approach used. I end the thesis with thoughts on future work.

Chapter 6

Discussion

We have explored three related models of local adaptation under eco-evolutionary dynamics in a heterogeneous environment with the ultimate aim of understanding how demographic and genetic stochasticity affect local adaptation.

Chapter 3 examines the foundational deterministic model in which an island receives maladapted migrants from a mainland source, whilst undergoing hard, directional selection and density regulation. We find that local adaptation depends sensitively on past events, the strength of selection relative to growth, the initial maladaptedness between locals and migrants and the number of loci. Crucially, we find that the coupling between population dynamics and population genetics gives rise to a second migration threshold. However, the significance of this threshold depends sensitively on trait architecture.

Chapter 4 explores the main premise of the thesis, namely that demographic and genetic stochasticity can have a striking effect on local adaptation. We build on the the deterministic model to include fluctuations in population size and allele frequencies (or trait average). Using the theory of diffusion processes, we model both a multilocus genetic model and a phenotypic model. Demographic drift causes extinction which impedes local adaptation. However, it also provides a means of escape: a maladapted population can shift to a better adapted state. Such transitions becomes less likely as the number of loci contributing to trait increase. We also consider phenotypic evolution under demographic and genetic drift. Our results depend on the assumption of a constant phenotypic variance. Under demographic drift, this assumption is essentially questionable, though it does make the dynamics more tractable.

The basic mainland-island model is extended in Chapter 5, where we consider an infinite island model with both deterministic and stochastic dynamics. In

short, multilocus evolution constrains conditions for generalism/polymorphism. Otherwise the single locus hard selection model is similar to the soft selection island models, in that polymorphism is possible between limited habitat frequencies. However, again, the genetic architecture of trait matters, since our results contrast strongly with trait models (though see Van Tienderen 1991). We have shown that demographic and genetic drift can reduce deme sizes, due to extinction and maladaptation. The effect of fluctuations is felt at very low migration rates with weak selection and strong drift. Then, extinction dominates and demes fix for one or other allele due to drift.

Issues: my work has focussed primarily on qualitative exploration of the models. I do not expect the results to be experimentally verifiable, especially because obtaining reliable estimates of biological parameters is notoriously difficult. Furthermore, changes in the details of the models will result in slightly different quantitative results. However, the overall qualitative results will remain largely unchanged, so long as the key elements of population regulation, directional selection and the coupling between population dynamics and evolution through migration remain unchanged.

Of course, as we have argued, should the trait architecture be considerably different or the form of selection, we expect different outcomes).

The primary motivation for this thesis was to consider local adaptation in the context of border populations subject to demographic and genetic drift. The question remains whether the continuous models are reasonable approximations of border populations. However, at the edge of the species' range, populations are small, fluctuate wildly and experience high migration (Hanski 1999). The diffusion approximation of Chapter 4 requires unrealistically low values for the intrinsic growth rate.

However, we believe that the continuous models are important in providing a qualitative understanding of the effect of demographic and genetic drift, as well as being easier to analyse than their discrete counterparts. In light of the 'irreducible uncertainty about mechanisms' (Coulson et al. 2006), perhaps a qualitative understanding is all we can hope for as models become more complex. The aim then should be to find generic and universal features, as I have tried to do. It is necessary to develop continuous models and discrete models in tandem, just as it is necessary to develop and compare genetic and trait models. There is a steady drive toward integrating ecological and evolutionary processes in a bid to introduce more realism into the highly idealised realm of theoretical work.

This thesis is a step in the right direction, and provides a possible foundation for future work. The models explored were deliberately kept relatively simple in order not to obfuscate the effect of demographic and genetic drift on local adaptation. However, there are many possible extensions and avenues for future work. A few are listed below.

6.1 Future work

Different types of selection: directional selection was modelled since it is the most prevalent form of selection in nature. It is also the simplest to model. However, different forms of selection, i.e. different fitness functions, can be examined quite easily in the diffusion framework of Chapter 4. For example, stabilising selection selects for intermediate values of the trait range, as opposed to the extreme values of the trait range (as in directional selection). This results in a fitness function that has a quadratic dependence on the allele frequencies (as opposed to a linear dependence on the allele frequencies due to directional selection) (Barton et al. 2007).

Fluctuating environment: the environment is clearly not constant and interacts with populations in a nontrivial way (Lande et al. 2009). This is an important potential theme for future work, to combine environmental interactions in models with demographic and genetic drift. This includes fluctuating selection, since contemporary evolution involves rapid changes in environmental conditions and temporal changes in selection due to anthropogenic disturbances (Hendry & Kinnison 1999). The potential dynamics of the stochastic model greatly simplifies the analysis of multivariate diffusion processes. However, the dependence on potential dynamics severely restricts the dynamics and hence how much realism can be incorporated into the diffusion formalism.

Linkage disequilibria: little work has been done on how population dynamics affects nonrandom associations between loci, let alone the effect of demographic stochasticity on linkage. We believe this is an important area of research that requires further exploration. One suggestion is to investigate a two-locus model which explicitly accounts for linkage disequilibria between loci. Multilocus evolution with linkage disequilibria and demographic drift can be studied with individual-based models that include realistic mating and density regulation.

Migration: the process of migration is not deterministic. Migrants can die along the way, so that migration and immigration rates between pairs of islands do not balance. Moreover, migration is often directed. Exploration of different forms of migration, which would invariably lead to population structures other than the infinite island model in Chapter 5.

Speciation: the stochastic infinite island model could be used to model quantitative traits under disruptive selection, as was done by (Barton & Rouhani 1993*a*,*b*). Models of Wright's shifting balance (Wright 1931, 1932) and speciation have not been studied with explicit population dynamics to the best of our knowledge.

In summary, for a qualitative understanding of how local adaptation is affected by the interaction between stochastic ecological and evolutionary processes, we believe the diffusion approximation is a sound approach. In particular, the stochastic models developed in this thesis provide a good foundation for exploration of more realistic assumptions, such as a changing environment.

Appendix A

Linear stability analysis

A.1 General linear stability analysis

The Jacobian matrix J_{ij} is given by $\partial(\partial x_i/\partial t)/\partial x_j$. Assuming we can linearise about the fixed points (Thompson & Stewart 1986), for the dynamics given in §3.2.1 (but using the scaled parameters), $\mathbf{J}$ has components

$$\begin{aligned} J_{1,1} &= \overline{r} - 2N \\ J_{i,1} &= -\frac{M}{N^2}(\hat{p}_i - p_i) && \text{for } i > 1, \\ J_{1,j} &= N\partial_{p_j}\overline{r} && \text{for } j > 1, \\ J_{i,j} &= \frac{(1-2p_i)}{2\eta}\partial_{p_i}\overline{r} - \frac{M}{N} && \text{if } i = j \text{ and } i, j > 1 \\ &= 0 && \text{otherwise,} \end{aligned} \tag{A-1}$$

where $\partial_{p_j}\overline{r} = 2s_j\eta$ In the general case this is an arrowhead matrix

$$\begin{pmatrix} J_{11} & J_{12} & J_{13} & J_{14} & J_{15} & \dots \\ J_{21} & J_{22} & 0 & 0 & 0 & \dots \\ J_{31} & 0 & J_{33} & 0 & 0 & \dots \\ J_{41} & 0 & 0 & J_{44} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{pmatrix}.$$

The eigenvalues of $\mathbf{J}$ are the roots of the characteristic polynomial $D = \det(\mathbf{J} - \lambda\mathbf{I}) = 0$, where $\mathbf{I}$ is a unit matrix of $n+1$ dimensions. Fortunately, the Laplacian

determinant expansion of an arrowhead matrix has the following tidy form

$$D = \prod_{i=1}^{n+1}(J_{i,i} - \lambda) - \sum_{j=2}^{n+1} J_{1,j}J_{j,1} \prod_{k\neq 1, k\neq j} (J_{k,k} - \lambda). \tag{A-2}$$

where $\prod_{k=h}^{h} = 1$. In principle it is possible to find exact expressions for λ. However, for $n > 3$ the expressions obtained are unwieldy and uninformative. On the other hand, numerical solution of D is fast.

If $M = 0$ or $s = 0$, $\mathbf{J}$ reduces to a triangular or diagonal matrix respectively, in which case the determinant is easily calculated as the product of diagonal elements. A nonlinear stability analysis is needed to determine the stability of the bifurcation points (Strogatz 2004).

A.1.1 Single locus stability analysis

The real parts of all eigenvalues of D must be negative for stability (Strogatz 2004). For the single locus case, the two eigenvalues of D are easily found:

$$\lambda_{+,-} = \frac{(J_{11} + J_{22}) \pm \sqrt{(J_{11} + J_{22})^2 - 4(J_{11}J_{22} - J_{12}J_{21})}}{2} \tag{A-3}$$

where

$$J_{11} = r - 2N, \qquad J_{12} = 2Ns,$$
$$J_{21} = \frac{M}{N^2}p, \qquad J_{22} = s(1-2p) - \frac{M}{N}.$$

Eigenvalue λ_+ is shown in Figure A.1. It determines the stability of the single locus states, since λ_- is always negative. For $M_l < M < M_h$, λ_+ is negative for both states, as expected. However this is not true for all s: when $s \leq 1$ the bistable regime disappears, see Fig. A.1b.

A.1.2 Graphical depiction of bistability

In Ch. 3, §3.3.4, we discuss diploid bistability. Emergence of the bistable regime and the changing stability of the steady states is depicted graphically in Fig. A.2. We define quartics $F_x = Q_x(0) + x$, where $x \in \{N, p\}$. Fixed points of F_x satisfy $F_x = x$, and these are just the steady state solutions (Strogatz 2004). For $M \ll M_l$, only the high state is stable. A qualitative change in the dynamics is signalled by a bifurcation at $M = M_l$. Bifurcation points often involve the

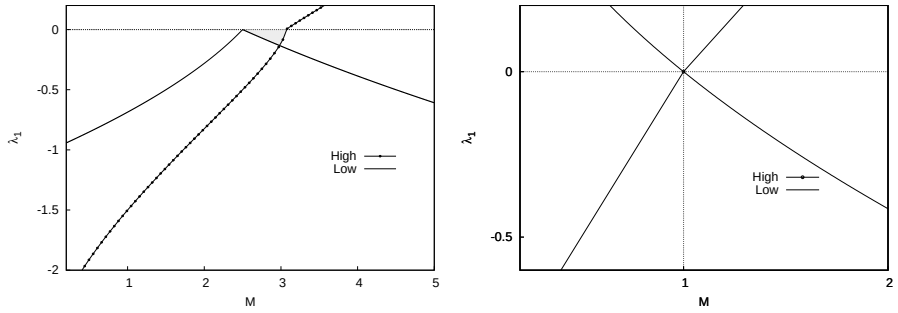


Figure A.1 *Eigenvalue λ_+ for the single locus states.(a) $s = 2.5$, bistability occurs when curves for both states are negative. (b) $s = 1$, the two states are never simultaneously negative.*

coalescence of two or more fixed points (Thompson & Stewart 1986). These are indicated by inverted triangles, see Figs. A.2a, A.2c and A.2g. The bifurcation in N is a saddle-node or fold bifurcation since the fixed point exists only for $M = M_l$. The bifurcation at $p = 0$ is transcritical—the fixed point exists for all values of M (Strogatz 2004). The presence of two stable nodes with an intermediate unstable node in both quartics is the signature of bistability, see Figs. A.2b and A.2f. The unstable state is $\{N_l^-, p_l\}$ (cf. 3.3.4). At $M = M_h$, where M_h is the upper bound of the bistability, $p_l = p_h$ and another qualitative change in the dynamics occurs; the bistability collapses (both bifurcations are saddle-nodes). Above rate M_h only the low state is stable.

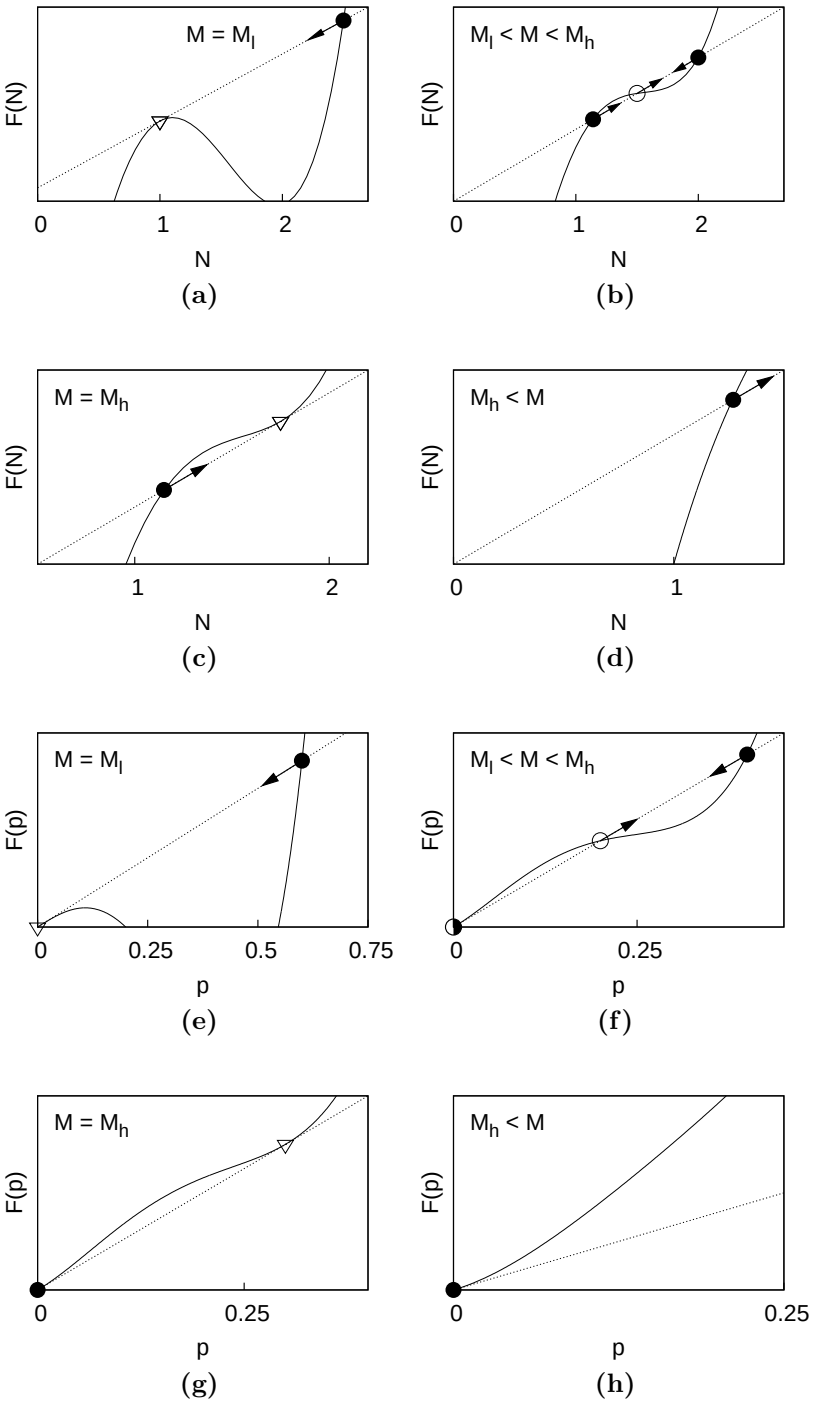


Figure A.2 *Fixed points of (a-d) $\dot{N} = 0$ and (e-h) $\dot{p} = 0$ for $s = 2.5$ and increasing M (left to right). From far left,* i*) $M = M_l$, the lower limit of the bistable regime. The only stable state (filled circle) state is the high state. Here the derivative of both curves at their respective maxima is 1.* ii*) $M_l < M < M_h$, the appearance of an unstable state (empty circle) in between two stable states is the hall mark of bistability.* iii*) $M = M_h$, bistability collapses.* iv*) $M_h < M$, there is only one stable state, the low state.*

Appendix B

Local peak approximation and the mean escape time

B.0.3 Local peak approximation

We outline the scheme for approximating the multilocus stationary distribution Ψ_n when its peaks are well-separated and sharply defined. The distribution can then be approximated as a sum of k peaks $\tilde{\Psi}_n \approx \frac{1}{Z_0}\sum_k \tilde{\Psi}_k$. Strictly, the peaks should be divided along the boundaries of their respective domains of attraction, the separatrices. Interior peaks and boundary peaks are treated separately. However the method of approximation is essentially the same in both cases; it is simply the Taylor expansion of potential U at peak k about a point $\mathbf{x}_k$.

We set the deterministic states as the location of the maxima i.e. we expand about the deterministic equilibria. This will be a good approximation to leading order in $O(\frac{1}{\zeta})$. It also avoids any problems arising from fluctuations taking the actual maximum of a boundary peak to unbiological values, such as negative p.

Interior peak

Assume Ψ_n has only interior peaks. An interior peak k that is clear of any boundaries can be approximated as a Gaussian $\tilde{\Psi}_{G,k}$ by applying the saddle-point approximation (Bender & Orszag 1991). Assuming the peak is narrowly spread about its maximum, $\mathbf{x}_k$, the potential is expanded about the peak maximum $\mathbf{x}_k$ so that

$$U(\mathbf{x}) \simeq U_k(\mathbf{x}) + \Big[\Big(\sum_i \delta_k x_i \partial_{x_i}\Big)U(\mathbf{x})\Big]_k + \frac{1}{2}\Big[\Big(\sum_i \delta_k x_i \partial_{x_i}\Big)^2 U(\mathbf{x})\Big]_k, \qquad \text{(B-1)}$$

where the subscript k indicates evaluation at $\mathbf{x}_k$ and $\delta_k x_i = x_i - x_{i,k}$, where $x_{i,k}$ is the i^{th} component of $\mathbf{x}_k$. The first order term is trivially zero since the expansion is taken about a stationary point.

Substituting (B-1) into Ψ_n at peak k gives the approximation

$$\tilde{\Psi}_{G,k} = \frac{1}{Z}\Phi_{n,k} e^{-\frac{1}{2}(\delta_k \mathbf{x})^T \mathbf{\Omega}_k (\delta_k \mathbf{x})} \tag{B-2}$$

where $\delta_k \mathbf{x} = \mathbf{x} - \mathbf{x}_k$ and $\Phi_{n,k} = \frac{e^{2\zeta U_k}}{N_k \prod_j v_{j,k}}$. Matrix $\mathbf{\Omega}$ has elements $\Omega_{ij} = -2\zeta \partial_{x_j} \partial_{x_i} U$ and is known as the precision or concentration matrix of a multivariate Gaussian. Its inverse is the covariance matrix $\mathbf{\Sigma}$, defined by $\Sigma_{ij} = \langle (x_i - \overline{x_i})(x_j - \overline{x_j}) \rangle$.

Normalisation of $\tilde{\Psi}_k$

To normalise $\tilde{\Psi}_{G,n}$ we utilise the fact that for small fluctuations the k^{th} peak probability is $P(k) = \int_{\sim k} \Psi_n \, \mathrm{d}\mathbf{x} = \frac{1}{Z} w_k$, where w_k is defined as the 'weight' of the peak and of course $\sum_k P(k) = 1$. Therefore $Z = \sum_k w_k$. Integration is across a small volume around the peak. Strictly speaking, this volume should be the domain of attraction of the equilibrium. However, dividing peaks along the unstable equilibria between peaks does not significantly affect the accuracy of the approximation. Using the expression for $P(k)$ together with $\int \Psi_n \, d\mathbf{x} = 1$, the final expression for $\tilde{\Psi}_{G,n}$ is

$$\tilde{\Psi}_{G,n} = \frac{1}{Z_0} \sum_k w_k e^{-\frac{1}{2}(\delta_k \mathbf{x})^T \mathbf{\Omega}_k (\delta_k \mathbf{x})}, \tag{B-3}$$

where weight $w_k = \Phi_{n,k} \sqrt{\frac{(2\pi)^{n+1}}{D_k}}$ and D_k is the determinant of $\mathbf{\Omega}$ evaluated at $\mathbf{x}_k$.

For the 'zero-width' approximation discussed in §4.3.2, we simply let the width of the Gaussian peaks in (B-3) go to zero so that,

$$\tilde{\Psi}_{0,n} \approx \frac{1}{Z_0} \sum_k w_k \delta(\mathbf{x} - \mathbf{x}_k), \tag{B-4}$$

where $w_k = \Phi_{n,k}$ and $\delta(\mathbf{x} - \mathbf{x}_k)$ is the Dirac delta function (not to be confused with $\delta_k \mathbf{x}$).

Average and maxima location

Under the local peak approximation, the average of each approximate Gaussian peak of $\tilde{\Psi}_{G,k}$ is just the maximum of that (normalised) peak. Now, the extrema of the exact distribution Ψ_n satisfy $\nabla \Psi_n = \mathbf{0}$,

$$0 = \dot{N} - \frac{1}{2\zeta} \tag{B-5a}$$

$$0 = \dot{p}_j - \frac{\partial_{p_j} v}{2N\zeta v}, \tag{B-5b}$$

where we have used the potential form of the deterministic dynamics to replace the derivatives of U. Assuming all loci are symmetric (and $\alpha = 1$), implicit solutions of the extrema are easily found to be

$$N^*_{\pm} = \frac{\overline{r} \pm \sqrt{(\overline{r})^2 + 2(2M - \frac{1}{\zeta})}}{2} \tag{B-6a}$$

$$p^*_{j,\pm} = \frac{(Ns - M + \frac{1}{2\zeta\eta}) \pm \sqrt{(Ns - M + \frac{1}{2\zeta\eta})^2 - 2Ns(2M\hat{p}_j - \frac{1}{2\zeta\eta})}}{2Ns}. \tag{B-6b}$$

In the limit of large ζ, the exact maxima can be approximated by an expansion in powers of ζ^{-1}, $\mathbf{x}_k \approx \mathbf{y}_{0,k} + \zeta^{-1}\mathbf{y}_{1,k} + \frac{1}{2!}\zeta^{-2}\mathbf{y}_{2,k} + \ldots$, where $\mathbf{y}_{0,k}$ is the k^{th} deterministic equilibrium. To zeroth order in ζ^{-1} then, the maxima of the Gaussian peaks $\tilde{\Psi}_{G,k}$ coincide with the deterministic equilibria. For all local peak approximations we set the maxima in this way. Also, dividing each peak along the deterministic unstable equilibria between peaks instead of the separatrices between domains of attraction will be accurate to leading order in ζ^{-1}.

Matrices $\mathbf{\Omega}$ and $\mathbf{\Sigma}$

Matrix $\Omega_{ij} = -2\zeta\partial_{x_j}\partial_{x_i}U$. In terms of the deterministic dynamics $\dot{x}_i$, $\partial_{x_i}U = \frac{\dot{x}_i}{B_i}$, where $B_i = B_{ii}$, the infinitesimal variance of x_i (see Eq. (4.4) in §4.2.1). Strictly speaking $\partial_{x_j}\partial_{x_i}U = \frac{1}{B_i}\partial_{x_j}\dot{x}_i + \dot{x}_i\partial_{x_j}\frac{1}{B_i}$ and must be evaluated at the maximum $\mathbf{x}_k$. If however we set the maximum of each peak to be the corresponding deterministic state $\mathbf{y}_k$, then $\partial_{x_j}\partial_{x_i}U = \frac{1}{B_i}\partial_{x_j}\dot{x}_i$, which is related to the Jacobian matrix $J_{ij} = \partial_{x_j}\partial_{x_i}\dot{x}_i$ (see App. A).

Matrix $\mathbf{\Omega}$ is positive definite and symmetric. Furthermore, it is an arrowhead matrix and for n (symmetric) loci it has the following the pleasing form

$$\begin{pmatrix} a & b & b & b & \dots \\ b & c & 0 & 0 & \dots \\ b & 0 & c & 0 & \dots \\ b & 0 & 0 & c & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

where

$$\begin{aligned} a &= 2\zeta\Big(1+\frac{M}{N^2}\Big), \\ b &= -4\zeta s\eta, \\ c &= \frac{M\zeta}{\eta v}\Big(\frac{1}{2\eta}+\frac{(\hat{p}-p)(1-2p)}{2\eta v}\Big), \end{aligned} \tag{B-7}$$

The covariance matrix is found by inverting the matrix $\mathbf{\Omega}$ and evaluating it at $\mathbf{x}_k$. Therefore the elements of matrix $\mathbf{\Sigma}$ are

$$\begin{aligned} \Sigma_{11} &= cQ_n, \\ \Sigma_{ij} &= -bQ_n && \text{for } i=1,\, j>1 \text{ or } j=1,\, i>1, \\ \Sigma_{ij} &= \frac{1}{c}\frac{Q_n}{Q_{n-1}} && \text{for } i=j,\, i,\, j>1, \\ \Sigma_{ij} &= \frac{b^2}{c}Q_n && \text{otherwise.} \end{aligned} \tag{B-8}$$

where $Q_n = \frac{1}{ac-nb^2}$. For a one-dimensional distribution, the inverse of the curvature of the peak would approximately be the variance of the peak. For higher dimensions, the curvature of a peak has no simple relation to the moments of the distribution. By determining $\mathbf{\Sigma}$ we can determine, for example, $\sigma^2_{x_i,k}$, the variance of x_i at peak k, which is simply Σ_{ii} evaluated at $\mathbf{x}_k$.

Stationary variance using $\tilde{\Psi}_k$

Using the Gaussian approximation $\tilde{\Psi}_{G,n}$ in (B-3), the stationary variance of x_i, can be calculated from the first two raw moments, $\sigma^2_{x_i} = \langle x_i^2\rangle - (\langle x_i\rangle)^2$. The average population size is

$$\langle x_i\rangle = \frac{1}{Z_0}\sum_k \Phi_{n,k} D_k x_{i,k} \tag{B-9}$$

where $x_{i,k}$ is the i^{th} component of $\mathbf{x}_k$. Using the result

$$\int x_i x_j e^{-\frac{1}{2}(\mathbf{x})^T \mathbf{A}(\mathbf{x})}\, d\mathbf{x} = (\mathbf{A}^{-1})_{ii} \sqrt{\frac{(2\pi)^d}{\det(\mathbf{A})}}, \tag{B-10}$$

for a d-dimensional Gaussian, the second raw moment is

$$\langle x_i^2 \rangle = \frac{1}{Z_0} \sqrt{\frac{(2\pi)^{n+1}}{D_k}} \Big((\mathbf{\Omega}_k^{-1})_{ij} + (x_{i,k})^2 \Big). \tag{B-11}$$

Putting (B-9) and (B-11) together gives $\sigma_{x_i}^2$.

Boundary peaks

For boundary peaks, essentially all we do is replace the Gaussian distribution in the above analysis with a Gamma distribution. However, there are some subtleties involved.

Consider a single boundary peak k of Ψ_n, for which $N \ll 1$ and $p_j \ll 1$. Discarding terms of order p^2 and N^2, the deterministic dynamics can be rewritten as

$$\dot{N} \approx \overline{r} N + M \tag{B-12a}$$

$$\dot{p}_j \approx s p_j + \frac{M}{N} (\hat{p}_j - p_j). \tag{B-12b}$$

With $B_{p_j} \approx \frac{p_j}{2\eta N}$ and $B_N = N$, the potential function is $U = N\overline{r} + M \ln(N) + 2M\eta \sum_j \hat{p}_j \ln(p_j) - p_j$. The potential is Taylor expanded about $\mathbf{x}_k$ as before, but this time we discard second order terms, such as N^2, p^2, Nsp,

$$\tilde{U}_k \approx U_k + M \ln\Big(\frac{N}{N_k}\Big) - |r_{min}| \delta_k N + 2M\eta \sum_j \hat{p}_j \ln\Big(\frac{p_j}{p_{j,k}}\Big) - 2M\eta \sum_j \delta_k p_j, \tag{B-13}$$

giving

$$\tilde{\Psi}_{\Gamma,k} = \frac{1}{Z_0} e^{-2\zeta |r_{min}| N} N^{2\zeta M - 1} \prod_j e^{-4M\zeta\eta p_j} p_j^{4M\zeta\eta\hat{p}_j - 1} \tag{B-14}$$

for the k^{th} boundary peak approximation (constant terms have been included in the normalisation, Z_0). Potential U separates into $n+1$ independent components when the coupling term between N and p_j, $e^{2N\zeta\overline{r}}$, is weak and can be set to 1

i.e. if $Ns\zeta p_j \ll 1$ (note that selection can be weak if n is large). For a number of such boundary peaks (after normalisation), this gives

$$\tilde{\Psi}_n \approx \frac{1}{\prod_l w_l} \sum_k h_N(2\zeta|\overline{r}_{min}|,\, 2M\zeta - 1) g_{N_k}(N_k; 2\zeta|\overline{r}_{min}|,\, 2M\zeta - 1) \times \tag{B-15}$$

$$\prod_j h_{p_j}(4M\zeta\eta,\, 4M\zeta\eta\hat{p}_j - 1) g_{p_{j,k}}(p_{j,k}; 4M\zeta\eta,\, 4M\zeta\eta\hat{p}_j - 1), \tag{B-16}$$

where $g(x; a, b) = e^{-ax}x^b$, $h(a, b) = a^{-b}\Gamma(b)$ and $w_l = h_N \prod_j h_{p_j}$, as given in §4.3.2 for a single locus.

B.0.4 Multivariate mean first passage time

Here we derive the mean exit time result, τ_k, from (4.37), which is essentially Talkner's multivariate mean exit time result, see Talkner 1987, Eq. 4.26. We follow the more accessible derivation in Hänggi et al. (1990, pp. 295-97).

Consider the first passage time $t(\mathbf{x})$ (FPT) out of the domain of attraction Λ, with boundary $\partial\Lambda$. The FPT satisfies

$$L_{BFP} t(\mathbf{x}) = -1 \tag{B-17}$$

in Λ, where

$$L_{BFP} = \sum_i A_i(\mathbf{x}) \partial_{x_i} + \frac{1}{2\zeta} \sum_i (\zeta B_i(\mathbf{x})) \partial^2_{x_i} \tag{B-18}$$

is the backward Fokker-Planck equation for the model, with the usual interpretation for $A_i(\mathbf{x})$ and $B_i(\mathbf{x})$. Noise parameter $1/2\zeta$ is equivalent to ϵ in Talkner (1987) and Hänggi et al. (1990). The mean first passage time $t(\mathbf{x})$ (MFPT) can be written as

$$t(\mathbf{x}) = t_{mfpt}(\mathbf{x}) g(\mathbf{x}), \tag{B-19}$$

where function $g(\mathbf{x})$ satisfies

$$g(\mathbf{x}) \approx 1 \qquad \text{inside the interior of } \Lambda, \tag{B-20}$$

$$= 0 \qquad \mathbf{x} \in \partial\Lambda, \tag{B-21}$$

and varies only within a small boundary layer at $\partial\Lambda$. Taking t^{-1}_{mfpt} as negligibly

small at the boundary, the form function satisfies $L_{BFP}(g(\mathbf{x})) \approx 0$. The boundary $\partial\Lambda$ is the separatrix of the deterministic drift field $\mathbf{A}(\mathbf{x})$. The MFPT, t_{mfpt}, can be expressed in terms of $g(\mathbf{x})$ and Ψ, a stationary solution of the forward Fokker-Plank equation,

$$L_{FFP} = -\sum_i \partial_{x_i} A_i(\mathbf{x}) + \frac{1}{2\zeta}\sum_i \partial^2_{x_i}(\zeta B_i(\mathbf{x})), \tag{B-22}$$

that obeys $L_{FFP}\Psi(\mathbf{x}) = 0$. By multiplying Eq. (B-17) by Ψ, integrating over domain Λ and invoking Gauss' theorem, we find,

$$t_{mfpt} = -\frac{2\zeta \int_\Lambda \Psi(\mathbf{x})\, d\mathbf{x}}{\sum_i \int_{\partial\Lambda}(\zeta B_i(\mathbf{x}))\Psi(\mathbf{x})\partial_{x_i} g(\mathbf{x})\, \mathrm{d}S_i}, \tag{B-23}$$

where $\mathrm{d}S$ is the oriented surface element on $\partial\Lambda$. Suppose the domain in question is the ecological niche (adaptive peak). Assuming low drift, the numerator will be dominated by peak density of the adaptive peak attractor. Therefore, to leading order in $1/\zeta$, $P(+) \equiv \int_\Lambda \Psi(\mathbf{x})\, d\mathbf{x}$, where $P(+)$ is the stationary probability of finding the population near the adaptive peak and can be approximated using the local Gaussian approximation discussed earlier, see Eq. B-2. The denominator in Eq. B-23 is dominated by the saddle point on the separatrix. The stationary probability distribution at the saddle point can be locally expanded about the saddle point, in much the same way as the local expansion of the adaptive peak. The key difference is that the matrix Ω evaluated at the saddle point will have one (and only one) negative eigenvalue, corresponding to the unstable direction at the saddle point due to the unstable state of the deterministic field.

We now only need to find the form function $g(\mathbf{x})$. We use the ansatz

$$g(\mathbf{x}) = \sqrt{\frac{4\zeta}{\pi}} \int_0^{k(\mathbf{x})} \exp(-z^2/\zeta)\, \mathrm{d}z, \tag{B-24}$$

where $k(\mathbf{x})$ is zero at the separatrix and increases as $\mathbf{x}$ leaves the separatrix. Substituting this into in $L_{BFP}(g(\mathbf{x})) \approx 0$, a first-order homogeneous differential equation for $k(\mathbf{x})$ is obtained. This is then solved by $k(\mathbf{x}) = \varsigma\left(\frac{B_{\varsigma\varsigma}}{\lambda_0}\right)^{-1/2}$, where ς is the unstable direction of the deterministic field at the saddle point and $B_{\varsigma\varsigma}$ is the diffusion matrix element in this direction. Using the results for $k(\mathbf{x})$, $g(\mathbf{x})$, and the local expansions of the probability distribution at the adaptive peak and the saddle point, we find

$$\sum_i \int_{\partial\Lambda} (\zeta B_i(\mathbf{x}))\Psi(\mathbf{x})\partial_{x_i} g(\mathbf{x})\, \mathrm{d}S_i \simeq -\lambda_0\left(\frac{\pi}{2\zeta}\right)^{-1} P(0) \tag{B-25}$$

where λ_0 is the positive eigenvalue at the saddle and $P(0)$ is the probability of finding the population at the saddle. This gives the final result

$$t_{mpft} = \frac{\pi}{\lambda_0} \frac{P(+)}{P(0)} \tag{B-26}$$

which leads to the expression for the rate in (4.38), if we identify $\tau_k = t_{mfpt}$.

References

Abramowitz, M. & Stegun, I. A., eds (1972), *Handbook of Mathematical Functions*, New York: Dover.

Aguileé, R., Claessen, D. & Lambert, A. (2009), 'Allele fixation in a dynamic metapopulation: founder effects vs refuge effects', *Theor. Pop. Biol.* **76**, 105–117.

Alleaume-Benharira, M., Pen, I. R. & Ronce, O. (2006), 'Geographic patterns of adaptation within a species' range: interactions between drift and gene flow', *J. Evol. Biol.* **19**, 203–219.

Anderson, W. W. (1971), 'Genetic equilibrium and population growth under density-regulated selection', *Am. Nat.* **105**, 489–498.

Andrewartha & Birch (1954), *The Distribution and Abundance of Animals*, Chicago: Chicago University Press.

Antonovics, J. (1976), 'The nature of limits to natural selection', *Ann. Missouri Bot. Gard.* **63**, 224–247.

Balakrishna, N., ed. (1991), *Handbook of the Logistic Distribution*, Marcek Dekker.

Barton, N. H. (1987), 'The probability of establishment of an advantaged mutant in a subdivided population', *Genet. Res.* **50**, 35–40.

Barton, N. H. (1989), 'The divergence of a polygenic system subject to stabilising selection, mutation and drift', *Genet. Res.* **54**, 59–77.

Barton, N. H. (1992), 'On the spread of new gene combinations in the third phase of Wright's shifting-balance', *Evolution* **46**, 551–557.

Barton, N. H. (1993), 'The probability of fixation of a favoured allele in a subdivided population', *Genet. Res.* **62**, 149–157.

Barton, N. H. (1999), 'Clines in polygenic traits', *Genet. Res.* **74**, 223–236.

Barton, N. H. (2001), Adaptation at the edge of the species' range, *in* J. Silvertown & J. Antonovics, eds, 'Integrating Ecology and Evolution in a Spatial Context', Oxford: Blackwell Science, pp. 365–392.

Barton, N. H., Briggs, D. E. G., Eisen, J. A., Goldstein, D. B. & Patel, N. H. (2007), *Evolution*, New York: Cold Spring Harbour Laboratory Press.

Barton, N. H. & Charlesworth, B. (1984), 'Genetic revolutions, founder effects, and speciation', *Ann. Rev. Ecol. Syst.* **15**, 133–164.

Barton, N. H. & Rouhani, S. (1987*a*), 'The frequency of shifts between alternative equilibria', *J. Theor. Biol.* **125**, 397–418.

Barton, N. H. & Rouhani, S. (1987*b*), 'Speciation and the 'shifting balance' in a continuous population', *Theor. Pop. Biol.* **31**, 465–492.

Barton, N. H. & Rouhani, S. (1993*a*), 'Adaptation and the 'shifting balance"', *Genet. Res.* **61**, 57–74.

Barton, N. H. & Rouhani, S. (1993*b*), 'Group selection and the 'shifting balance"', *Genet. Res.* **61**, 127–135.

Barton, N. H. & Turelli, M. (1989), 'Evolutionary quantitative genetics: how little do we know?', *Annu. Rev. Genet.* **23**, 337–370.

Barton, N. H. & Whitlock, M. C. (1997), The evolution of metapopulations, *in* I. Hanski & M. E. Gilpin, eds, 'Metapopulation Biology: Ecology, Genetics and Evolution', San Diego: Academic Press, pp. 183–210.

Bender, C. M. & Orszag, S. A. (1991), *Advanvced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbations Theory*, New York: Springer-Verlag.

Blondel, J., Perret, P., Maistre, M. & Dias, P. C. (1992), 'Do harlequin mediterranean environments function as source sink for blue tits (*Parus caerueus*)?', *Landscape Ecol.* **6**, 213–219.

Blows, M. W. & Hoffman, A. A. (2005), 'A reassessment of genetic limits to evolutionary change', *Ecology* **86**, 1371–1384.

Bodmer, W. F. & Cavalli-Sforza, L. L. (1968), 'A magiration matrix model for the study of random genetic drift', *Genetics* **59**, 565–592.

Boilley, D., Juardo, B. & Schmitt, C. (2004), 'Simple relations between the mean passage times and kramers' stationary rate', *Phys. Rev. E* **70**, 056219.

Bradshaw, A. D. (1991), 'Genostasis and the limits to evolution', *Phil. Trans. Roy. Soc. B* **333**, 289–305.

Brassil, C. E. (2001), 'Mean time to extinction of a metapopulation with an allee effect', *Ecol. Model.* **143**, 9–16.

Bridle, J. R., Polechová, J., Kawata, M. & Butlin, R. K. (2010), 'Why is adaptation prevented at ecological margins? new insights from individual-based simulations', *Ecol. Lett.* **13**(4), 485–494.

Bridle, J. R. & Vines, T. H. (2006), 'Limits to evolution at range margins: when and why does adaptation fail?', *Trends Ecol. Evol.* **22**, 140–147.

Brown, J. H. (1995), 'Spatial variation in abundance', *Ecology* **76**, 2028–2043.

Brown, J. H. & Kodric-Brown, A. (1977), 'Turnover rates in insular biogeography: effect of immigration on extinction', *Ecology* **58**, 445–449.

Brown, J. H. & Lomolino, M. V. (1998), *Biogeography*, 2nd edn, Sunderland, MA: Sinauer.

Brown, J. S. & Pavlovic, N. B. (1992), 'Evolution in heterogeneous environments: effects of migration on habitat specialization', *Evol. Ecol.* **5**, 360–382.

Bulmer, M. G. (1972), 'Multiple niche polymorphism', *Am. Nat.* **106**, 254–257.

Bulmer, M. G. (1980), *The Mathematical Theory of Quantitative Genetics*, Oxford: Oxford University Press.

Bürger, R. & Lynch, M. (1995), 'Evolution and extinction in a changeing environment: a quantitative-genetic analysis', *Evolution* **49**(1), 151–163.

Burton, O. J. & Travis, J. M. J. (2008), 'The frequency of fitness peak shifts is increased at expanding range margins due to mutation surfing', *Genetics* **179**, 941–950.

Cain, A. J. & Sheppard, P. M. (1950), 'Selection in the polymorphic *Capaea Nemoralis*', *Heredity* **4**, 275–294.

Camin, J. H. & Ehrlich, P. R. (1958), 'Natural selection in water snakes (*Natrix sipedon* l.) on islands in lake erir', *Evolution* **12**, 504–511.

Caroli, B., Caroli, C., Roulet, B. & Gouyet, J. F. (1980), 'A wkb treatment of diffusion in a multidimensional bistable potential', *J. Stat. Phys.* **22**, 515–536.

Carroll, S. P., Hendry, A. P., Reznick, D. N. & Fox, C. W. (2007), 'Evolution on ecological time-scales', *Funct. Ecol.* **21**, 387–393.

Case, T. J. & Taper, M. L. (2000), 'Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders', *Am. Nat.* **155**, 583–605.

Charlesworth, B. (1971), 'Selection in density-regulated populations', *Ecology* **52**, 469–474.

Charlesworth, B., Charlesworth, D. & Barton, N. H. (2003), 'The effects of genetic and geographic structure on neutral variation', *Ann. Rev. Ecol. Evol. Syst.* **34**, 99–125.

Christiansen, F. B. (1974), 'Sufficient conditions for protected polymorphism in a subdivided population.', *Am. Nat.* **108**, 157–164.

Christiansen, F. B. (1975), 'Hard and soft selection in a subdivided population.', *Am. Nat.* **109**, 11–16.

Christiansen, F. B. (2004), Density-dependent selection, *in* R. S. Singh & M. K. Uyenoyama, eds, 'The Evolution of Population Biology', Cambirdge: Cambridge University Press.

Christiansen, F. B. & Feldman, M. W. (1975), 'Subdivided populations: a review of the one- and two-locus deterministic theory', *Theor. Pop. Biol.* **7**, 13–38.

Coltman, D. W. (2005), 'Differentiation by dispersal', *Nature* **433**, 23–24.

Coulson, T., Benton, T. G., Lundberg, P., Dall, S. R. X. & Kendall, B. E. (2006), 'Putting evolutionary biology back in the ecological theatre: a demographic framework for mapping genes to communities', *Evol. Ecol. Res.* **8**, 1155–1171.

Cox, D. R. & Miller, H. D. (1980), *The Theory of Stochastic Processes*, London: Chapman and Hall.

Coyne, J. A., Barton, N. H. & Turelli, M. (1997), 'Perspective: a critique of Wright's shifting balance theory of evolution', *Evolution* **51**, 643–671.

Crow, J. F., Engels, W. R. & Denniston, C. (1990), 'Phase three of Wright's shifting-balance theory', *Evolution* **44**, 233–247.

Crow, J. F. & Kimura, M. (1970), *An Introduction to Population Genetics Theory*, New York: Harper and Row.

Crow, J. F. & Maruyama, T. (1971), 'The number of neutral alleles maintained in a finite, geographically structured population', *Theor. Pop. Biol.* **2**, 437–453.

Deakin, M. A. B. (1966), 'Sufficient conditions for genetic polymorphism', *Am. Nat.* **100**, 690–692.

Deakin, M. A. B. (1968), 'Genetic polymorphism in a subdivided population', *Aust. J. Biol. Sci.* **21**, 165–168.

Dhondt, A. A., Adriaensen, F., Matthysen, E. & Kempenaers, B. (1990), 'Non-adaptive clutch sizes in tits', *Nature* **348**, 723–725.

Dias, P. C. (1996), 'Sources and sinks in population biology', *Trends Ecol. Evol.* **11**, 326–330.

Donelly, P. & Weber, N. (1985), 'The Wright-Fisher model with temporally varying selection and population size', *J. Math. Biol.* **22**, 21–29.

Ehrlich, P. R. & Raven, P. H. (1969), 'Differentiation of populations', *Science* **165**, 1228–1232.

Endler, J. A. (1973), 'Gene flow and population differentiation', *Science* **179**, 243–250.

Endler, J. A. (1977), *Geographic Variation, Speciation and Clines*, 1st edn, Princeton, NJ: Prinction University Press.

Engen, S., Lande, R. & S æther, B.-E. (2009), 'Fixation probability of beneficial mutations in a fluctuating population', *Genet. Res.* **91**, 73–82.

Ewens, W. J. (1967), 'The probability of survival of a new mutant in a fluctuating environment', *Heredity* **22**, 438–44.

Ewens, W. J. (1979), *Mathematical Population Genetics*, 1st edn, Berlin: Springer-Verlag.

Fagane, W. F., Lynch, H. J. & Noon, B. R. (2010), 'Pitfalls and challenges of estimating population growth rate from empirical data: consequences for allometric scaling relations', *Oikos* **119**, 455–464.

Falconer, D. S. & MacKay, T. F. C. (1996), *Introduction to Quantitative Genetics*, 3rd edn, Reading, MA: Addison-Wesley.

Feller, W. (1960), *An Introduction to Probability Theory and Its Application*, Vol. 1, New York: John Wiley and Sons.

Felsenstein, J. (1976), 'The theoretical population genetics of variable selection and migration', *Ann. Rev. Genet.* **10**, 253–280.

Filin, I., Holt, R. D. & Barfield, M. (2008), 'The relation of density regulation to habitat specialization, evolution of a species' range, and the dynamics of biological invasions', *Am. Nat.* **172**, 233–247.

Fisher, R. A. (1918), 'The correlation between relatives under the supposition of mendelian inheritance', *Trans. Roy. Soc. Edinburgh* **52**, 399–433.

Fisher, R. A. (1937), 'The wave of advance of an advantageous gene', *Ann. Eugen.* **7**, 355–369.

Fisher, R. A. (1950), 'Gene frequencies in a cline determined by selection and diffusion', *Biometrics* **6**, 353–61.

Fisher, R. A. (1958), *The Genetical Theory of Natural Selection*, 2nd edn, New York: Dover.

Fortin, W. C., Keitt, T. H., Maurer, B. A., Taper, M. L., Kauffman, D. M. & Blackburn, T. M. (2005), 'Species' geographic ranges and distributional limits: pattern analysis and statistical issues', *Oikos* **108**, 7–17.

Frankham, R. & Ralls, K. (1998), 'Inbreeding leads to extinction', *Nature* **392**, 441–442.

Gabriel, W. & Bürger, R. (1992), 'Survival of small populations under demographic stochasticity', *Theor. Pop. Biol.* **41**, 44–71.

Garcías-Ramos, G. & Kirkpatrick, M. (1997), 'Genetic models of adaptation and gene flow in peripheral populations', *Evolution* **51**, 21–28.

Garcías-Ramos, G. & Rodríguez, D. (2002), 'Evolutionary speed of species invasion', *Evolution* **56**, 661–668.

Gardiner, C. W. (1985), *Handbook of Stochastic Methods For Physics, Chemistry and the Natural Sciences*, 2nd edn, Springer-Verlag Berlin.

Gillespie, J. H. (1974), 'Polymorphism in patchy environments', *Am. Nat.* **108**, 145–151.

Gillespie, J. H. (1975), 'The role of migration in the genetic structure of populations in temporarily and spatially varying environments. i. conditions for polymorphism', *Am. Nat.* **109**, 127–135.

Gliddon, C. & Strobeck, C. (1975), 'Necessary and sufficient conditions for multiple-niche polymorphism in haploids', *Am. Nat.* **109**, 233–235.

Gomulkiewicz, R. & Holt, R. D. (1995), 'When does evolution by natural selection prevent extinction?', *Evolution* **49**, 201–207.

Gomulkiewicz, R., Holt, R. D. & Barfield, M. (1999), 'The effects of density dependence and immigration on local adaptation and niche evolution in a black hole sink environment', *Theor. Pop. Biol.* **55**, 283–296.

Goodnight, C. J. (2005), 'Peak shifts in large populations', *Heredity* **96**, 5–6.

Grant, P. R. & Grant, B. R. (2006), 'Evolution of character displacement in darwin's finches', *Science* **313**, 224–226.

Grasman, J. & Ludwig, D. (1983), 'The accuracy of the diffusion approximation to the expected time to extinction for some discrete stochastic processes', *J. Appl. Prob.* **20**, 305–321.

Hadany, L. (2002), 'Adaptive peak shifts in a heterogeneous environment', *Theor. Pop. Biol.* **63**, 41–51.

Haldane, J. B. S. (1931), 'A mathematical theory of natural selection. vi. isolation', *Trans. Camb. Phil. Soc.* **26**, 220–230.

Haldane, J. B. S. (1932), *The Causes of Evolution*, London: Longmans, Green and Co.

Haldane, J. B. S. (1956), 'The relation between density regulation and natural selection', *Proc. Roy. Soc. Lond. B* **145**, 306–308.

Haldane, J. B. S. (1958), 'Theory of a cline', *J. Genetics* **48**, 277–284.

Halley, J. M. & Iwasa, Y. (1998), 'Extinction rate of a population under both demographic and snvironmental stochasticity.', *Theor. Pop. Biol.* **53**, 1–15.

Hänggi, P. (1986), 'Escape from a metastable state', *J. Stat. Phys.* **42**, 105–148.

Hänggi, P., Talkner, P. & Borkovec, M. (1990), 'Reaction-rate theory: fifty years after Kramers', *Rev. Mod. Phys.* **62**, 251–341.

Hanski, I., ed. (1999), *Metapopulations Ecology*, Oxford: Oxford University Press.

Hanski, I. & Gilpin, M. E., eds (1997), *Metapopulations Biology: Ecology, Genetics and Evolution*, San Diego: Academic Press.

Harrison, S. & Taylor, A. D. (1997), Empirical evidence for metapopulation dynamics, *in* I. Hanski & M. E. Gilpin, eds, 'Metapopulations Biology: Ecology, Genetics and Evolution', San Diego: Academic Press.

Hedrick, P. W. (1986), 'Genetic polymorphism in heterogeneous environments: a decade later', *Ann. Rev. Ecol. Syst.* **17**, 535–566.

Hedrick, P. W. (2004), Conservation biology: the impact of population biology and a current perspective, *in* R. S. Singh & M. K. Uyenoyama, eds, 'The Evolution of Population Biology', Cambridge: Cambridge University Press, pp. 347–365.

Hedrick, P. W. (2006), 'Genetic polymorphism in heterogeneous environments: the age of genomics', *Ann. Rev. Ecol. Evol. Syst.* **37**, 67–93.

Hedrick, P. W., Ginevan, M. E. & Ewing, E. P. (1976), 'Genetic polymorphism in heterogeneous environments', *Ann. Rev. Ecol. Syst.* **7**, 1–32.

Hedrick, P. W., Lacy, R. C., Allendorf, F. W. & Soulé, M. E. (1996), 'Directions in conservation biology: comments on Caughley', *Conserv. Biol.* **10**, 11312–1320.

Hendry, A. P. & Kinnison, M. T. (1999), 'The pace of modern life: measuring rates of contemporary microevolution', *Evolution* **53**, 1637–1653.

Henle, K., Sarre, S. & Wiegand, K. (2004), 'The role of density regulation in extinction processes and population viability analysis', *Biodiv. Conserv.* **13**, 9–52.

Heyde, C. C. & Seneta, E. (1975), 'The genetic balance between random sampling and random population size', *J. Math. Biol.* **1**, 317–320.

Hill, M. F., Hastings, A. & Botsford, L. W. (2002), 'The effects of small dispersal rates on extinction times in structured metapopulation models', *Am. Nat.* **160**, 389–402.

Hoekstra, R. F., Bijlsma, R. & Dolman, A. J. (1985), 'Polymorphism from environmental heterogeneity: models are only robust if the heterozygote is close in fitness to the favoured homozygote in each environment', *Genet. Res.* **45**, 299–314.

Holt, R. D. (1983*a*), Immigration and the dynamics of peripheral populations, *in* A. G. J. Rhodin & K. Miyata, eds, 'Advances in Herpetology and Evolutionary Biology', Cambridge, MA: Harvard Univeristy, Museum of Comparative Zoology, pp. 680–694.

Holt, R. D. (1983*b*), Models for peripheral populations: the role immigration, *in* H. I. Freedman & C. Strobeck, eds, 'Population Biology: Lecture Notes in Biomathematics, Vol. 52', Berlin: Springer-Verlag, pp. 25–32.

Holt, R. D. (1996*a*), 'Adaptive evolution in source-sink environements: direct and indirect effects of density-dependence on niche evolution', *Oikos* **75**, 182–192.

Holt, R. D. (1996*b*), 'Demographic constraints in evolution: towards unifying the evolutionary theories of senescence and niche conservatism', *Evol. Ecol.* **10**, 1–11.

Holt, R. D. (1997), 'On the evolutionary stability of sink populations', *Evol. Ecol.* **11**, 723–731.

Holt, R. D. (2009), 'Bringing the Hutchinson niche into the 21st century: ecological and evolutionary perspectives', *Proc. Nat. Acad. Sci.* **106**, 19659–19665.

Holt, R. D., Barfield, M. & Gomulkiewicz, R. (2005), Theories of niche conservatism and evolution. Could exotic species be potential tests?, *in* D. F. Sax, J. J. Stachowicz & S. D. Gaines, eds, 'Species Invasions: Insights into Evolution, Ecology and Biogeography', Sunderland, MA: Sinauer, pp. 259–290.

Holt, R. D. & Gaines, M. S. (1992), 'Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches', *Evol. Ecol.* **6**, 433–447.

Holt, R. D. & Gomulkiewicz, R. (1997), 'How does immigration influence local adaptation? a reexamination of a familiar paradigm', *Am. Nat.* **149**, 563–572.

Holt, R. D., Gomulkiewicz, R. & Barfield, M. (2003), 'The phenomenology of niche evolution via quantitative traits in a 'black-hole' sink', *Proc. Roy. Soc. Lond. B.* **270**, 215–224.

Holt, R. & Gomulkiewicz, R. (2002), Conservation implications of niche conservatism and evolution in heterogeneous environments, *in* U. Dieckmann, R. Ferrière & D. Couvet, eds, 'Evolutionary Conservation Biology', Cambridge: Cambridge University Press, pp. 244–264.

Hutchinson, G. E. (1958), 'Concluding remarks', *Cold Spring Harbour Symp. Quant. Biol.* **22**, 415–427.

Kaitala, V., Ranta, E. & Stenseth, N. C. (2006), 'Genetic structuring in fluctuating populations.', *Ecol. Infor.* **1**, 343–348.

Karlin, S. (1968), 'Rates of approach to homozygosity for finite stochastic models with variable population size', *Am. Nat.* **102**, 443–455.

Karlin, S. (1976), Population subdivision and selection-migration interaction, *in* S. Karlin & E. Nevo, eds, 'Population Genetics in Ecology', New York: Academic Press, pp. 617–657.

Karlin, S. & Taylor, H. M. (1975*a*), *A First Course in Stochastic Processes*, 2nd edn, New York: Academic Press.

Karlin, S. & Taylor, H. M. (1975*b*), *A Second Course in Stochastic Processes*, 2nd edn, New York: Academic Press.

Kawecki, T. J. (1995), 'Demography of source-sink populations and the evolution of ecological niches', *Evol. Ecol.* **9**, 32–44.

Kawecki, T. J. (2000), 'Adaptation to marginal habitats: contrasting influence of dispersal on the fate of rare alleles with small and large effects', *Proc. Roy. Soc. Lond. B* **267**, 1315–1320.

Kawecki, T. J. (2008), 'Adaptation to marginal habitats', *Ann. Rev. Ecol. Evol. Syst.* **39**, 321–342.

Kawecki, T. J. & Ebert, D. (2004), 'Conceptual issues in local adapatation', *Ecol. Lett.* **7**, 1225–1241.

Kawecki, T. J. & Holt, R. D. (2002), 'Evolutionary consequences of asymmetric dispersal rates', *Am. Nat.* **106**, 333–347.

Kawecki, T. J. & Stearns, S. C. (1993), 'The evolution of life histories in spatially heterogeneous environments: optimal reaction norms revisited', *Ecol. Evol.* **7**, 155–174.

Keitt, T. H., Lewis, M. A., & Holt, R. D. (2000), 'Allee effects, invasion pinning, and species' borders', *Am. Nat.* **157**, 203–216.

Kettlewell, H. B. D. (1973), *The Evolution of Melanism*, Oxford: Oxford University Press.

Kimura, M. (1962), 'On the probability of fixation of mutant genes in a population', *Genetics* **47**, 713–719.

Kimura, M. & Ohta, T. (1968), 'The average number of generations until fixation of a mutant gene in a finite population', *Genetics* **61**, 763–771.

Kimura, M. & Ohta, T. (1974), 'Probability of gene fixation in an expanding finite population', *Proc. Natl. Acad. Sci.* **71**, 3377–3379.

Kimura, M. & Weiss, G. H. (1964), 'The stepping stone model of population structure and the decrease of genetic correlation with distance', *Genetics* **49**, 561–576.

King, R. B. & Lawson, R. (1995), 'Color-pattern variation in lake erie water snakes: the role of gene flow.', *Evolution* **49**, 885–896.

Kingslover, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., Hoang, A., Gilbert, P. & Beerli, P. (2001), 'The strength of phenotypic selection in natural populations', *Am. Nat.* **157**, 245–261.

Kinnison, M. T. & Hendry, A. P. (2001), 'The pace of modern life II: from rates of contemporary microevolution to pattern and process', *Genetica* **112-113**, 145–164.

Kinnison, M. T., Hendry, A. P. & Stockwell, C. A. (2007), 'Contemporary evolution meets conservation biology II: impediments to integration and application', *Ecol. Res.* **22**, 947–954.

Kinnison, M. T. & Jr., N. G. H. (2007), 'Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence', *Funct. Ecol.* **21**, 444–454.

Kirkpatrick, M. & Barton, N. H. (1997), 'Evolution of a species' range', *Am. Nat.* **150**, 1–23.

Lambert, A. (2006), 'Probability of fixation under weak selection: a branching process unifying approach', *Theor. Pop. Biol.* **69**, 419–441.

Lande, R. (1976), 'Natural selection and random drift in phenotypic evolution', *Evolution* **30**, 314–334.

Lande, R. (1979), 'Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry', *Evolution* **34**, 402–416.

Lande, R. (1984), 'The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization', *Heredity* **54**, 323–332.

Lande, R. (1985*a*), 'Expected time for random genetic drift of a population between stable phenotypic states', *Proc. Natl. Acad. Sci.* **82**, 7641–7645.

Lande, R. (1985*b*), 'The fixation of chromosomal rearrangements in assubdivided population with local extinction and colonization', *Heredity* **54**, 323–332.

Lande, R. (1985*c*), 'Genetic variation and phenotypic evolution during allopatric speciation', *Am. Nat.* **116**(4), 463–479.

Lande, R. (1986), 'The dynamics of peak shifts and the pattern of morphological evolution', *Paleobiology* **12**, 343–354.

Lande, R. (1988), 'Genetics and demography in biological conservation', *Science* **241**, 1455–1460.

Lande, R. (1992), 'Neutral theory of quantitative genetic variance in an island model with local extinction and colonization', *Evolution* **46**, 381–389.

Lande, R. (1993), 'Risks of population extinction from demographic and environmental stochasticity and random catastrophes', *Am. Nat.* **142**, 911–927.

Lande, R., Engen, S. & S æther, B. (1998), 'Extinction times in finite metapopulation models with stochastic local dynamics', *Oikos* **83**, 383–389.

Lande, R., Engen, S. & S æther, B.-E. (2009), 'An evolutionary maximum principle for density-dependent population dynamics in a fluctuating environment', *Phil. Soc. Trans. Roy. Soc. B.* **364**, 1511–1518.

Lande, R., Engen, S. & Saether, B. (2003), *Stochastic Population Dynamics in Ecology and Conservation*, 1st edn, Oxford: Oxford University Press.

Lawton, J. H. (1993), 'Range, population abundance and conservation', *Trends Ecol. Evol.* **8**, 353–359.

Lawton, J. H. & Woodroffe, G. L. (1991), 'Habitat and distribution of water voles: why are there gaps in a species' range?', *J. Anim. Ecol.* **60**, 79–91.

Leigh, E. (1981), 'The average lifetime of a population in a varying environment', *J. Theor. Biol.* **90**, 213–239.

Lenormand, T. (2002), 'Gene flow and the lmits to natural selection', *Trends Ecol. Evol.* **17**, 183–189.

Levene, H. (1953), 'Genetic equilibrium when more than one ecological niche is available', *Am. Nat.* **87**, 331–333.

Levin, S. (1978), On the evolution of ecological parameters, *in* P. F. Brussard, ed., 'Ecological Genetics: The Interface', New York: Springer-Verlag, pp. 3–26.

Lewontin, R. C. (2004), Building a science of population biology, *in* R. S. Singh & M. K. Uyenoyama, eds, 'The Evolution of Population Biology', Cambirdge University Press, Cambridge.

Ludwig, D. (1976), 'A singular perturbation problem in the theory of population extinction', *SIAM Am. Math. Soc. Proc.* **10**, 87–104.

Ludwig, D. (1996), 'The distribution of population survival times', *Am. Nat.* **147**, 506–526.

Lynch, M. & Gabriel, W. (1990), 'Mutational load and survival of small populations', *Evolution* **44**, 1725–1737.

Lynch, M., Gabriel, W. & Wood, A. M. (1991), 'Adaptive and demographic responses of plankton populations to environmental change', *Limnol. Ocean.* **36**, 1301–1312.

Lynch, M. & Lande, R. (1993), Evolution and extinction in response to environmental change, *in* P. M. Kareiva, J. G. Kingslover & R. B. Huey, eds, 'Biotic Interactions and Global Change.', Sunderland, MA: Sinauer, pp. 234–250.

MacArthur, R. H. (1962), 'Some generalized theorems of natural selection', *Proc. Natl. Acad. Sci.* **48**, 1893–7.

MacArthur, R. H. & Wilson, E. O. (1967), *The Theory of Island Biogeography*, Princeton: Princeton University.

Macnair, M. R. (1981), Tolerance of higher plants to toxic materials, *in* J. A. Bishop & L. M. Cook, eds, 'Genetic Consequences of Man Made Change', 1st edn, New York: Academic Press, pp. 177–207.

MacNair, M. R. (1987), 'Heavy metal tolerance in plants: A model evolutionary system', *Trends Ecol. Evol.* **2**, 354–359.

Mangel, M. & Tier, C. (1993*a*), 'Dynamics of metapopulations with demographic stochasticity and environmental catastrophes', *Theor. Pop. Biol.* **44**, 1–31.

Mangel, M. & Tier, C. (1993*b*), 'A simple direct method for finding persistence times of populations and applications to conservation biology', *Proc. Natl. Acad. Sci.* **90**, 1083–1086.

Maruyama, T. (1971), 'Analysis of population structure. II two dimensional stepping stone models of finite length and other geographically structured populations', *Ann. Human. Genet.* **35**, 179–196.

Maruyama, T. (1972), 'Rate of decrease of genetic variability in two-dimensional continuous population of finite size', *Genetics* **70**, 639–651.

Maruyama, T. (1974), 'A simple proof that certain quantities are independent of the geographic structure of a population', *Theor. Pop. Biol.* **5**, 148–154.

Maruyama, T. & Kimura, M. (1980), 'Genetic variability and effective population size when local extinction and recolonization of subpopulations are frequent', *Proc. Natl. Acad. Sci.* **77**, 6710–6714.

May, R. M. (1974), 'Biological populations with non-overlapping generations: stable points, stable cycles and chaos', *Science* **186**, 645–647.

May, R. M. (1976), 'Simple mathematical models with very complicated dynamics', *Nature* **167**, 459–467.

May, R. M. & Oster, G. F. (1976), 'Bifurcations and dynamic complexity in simple ecological models', *Am. Nat.* **110**, 573–599.

Maynard Smith, J. (1966), 'Sympatric speciation', *Am. Nat.* **100**(916), 637–650.

Maynard Smith, J. (1970), 'Genetic polymorphism in a varied envirionment.', *Am. Nat.* **104**, 487–490.

Mayr, E. (1963), *Animal Species and Evolution*, Cambridge, MA: Harvard University Press.

McCauley, D. E. (1991), 'Genetic consequences of local population extinction and recolonization', *Trends Ecol. Evol.* **6**, 5–8.

McNeilly, T. (1968), 'Evolution in closely adjacent plant poplations III. *Agrostis tenuis* on a small copper mine', *Heredity* **23**, 99–108.

Moore, J. & Hendry, A. P. (2009), 'Can gene flow have negative demographic consequences? mixed evidence from stream threespine stickleback', *Phil. Trans. Roy. Soc. B* **463**, 1533–1542.

Moran, P. (1962), *Statistical Process in Evolutionary Theory*, Oxford: Oxford University Press.

Murray, J. D. (1989), *Mathematical Biology*, Berlin: Springer-Verlag.

Nagylaki, T. (1975), 'Conditions for the existence of clines', *Genetics* **80**, 595–615.

Newman, C. M., Cohen, J. E. & Kipnis, C. (1985), 'Neo-darwinian evolution impllies punctuated equilibria', *Nature* **315**, 400–401.

Nieminen, M., Singer, M. C., Fortelius, W., Schöps, K. & Hanski, I. (2001), 'Experimental confirmation that inbreeding depression increases extinction risk in butterfly populations.', *Am. Nat.* **157**, 237–244.

Oakeshott, J. G., Gibson, J. B., Anderson, P. R., Knibb, W. R., Anderson, D. G. & Chambers, G. K. (1982), 'Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila Melanogaster* on different continents', *Evolution* **36**, 86–96.

O'Grady, J. J., Brook, B. W., Reed, D. H., Ballou, J. D., Tonkyn, D. W. & Frankham, R. (2006), 'Realistic level of inbreeding depression strongly affect extinction risk in wild populations', *Biol. Conserv.* **133**, 42–51.

Otto, S. P. & Whitlock, M. C. (1997), 'The probability of fixation in populations of changing size', *Genetics* **146**, 723–733.

Pannell, J. R. (2003), 'Coalescent in a metapopulation with recurrent local extinction and recolonization', *Evolution* **57**(5), 949–961.

Parsons, T. L. & Quince, C. (2007*a*), 'Fixation in haploid populations exhibiting density dependence I: The non-neutral case', *Theor. Pop. Biol.* **72**, 121–135.

Parsons, T. L. & Quince, C. (2007*b*), 'Fixation of haploid populations exhibiting density dependence II: The quasi-neutral case', *Theor. Pop. Biol.* **72**, 468–479.

Parsons, T. L., Quince, C. & Plotkin, J. B. (2008), 'Absorption and fixation times for neutral and quasi-neutral populations with density dependence', *Theor. Pop. Biol.* **74**, 302–310.

Pease, C. M., Lande, R. & Bull, J. J. (1989), 'A model of population growth, dispersal and evolution in a changing environment', *Ecology* **70**, 1657–1664.

Peck, S. L., Ellner & Gould (1998), 'A spatially explicit stochastic model demonstrates the feasibility of Wright's shifting balance theory', *Evolution* **52**, 1834–1839.

Pelletier, F., Garant, D. & Hendry, A. P. (2009), 'Eco-evolutionary dynamics', *Phil. Trans. Roy. Soc. B* **364**, 1483–1489.

Polechová, J., Barton, N. H. & Marion, G. (2009), 'Species' range: adaptation in space and time', *Am. Nat.* **174**, E186–E204.

Pollak, E. (1972), 'Some effects of two types of migration on the survival of a gene', *Biometrics* **28**, 385–400.

Pollak, E. (1974), 'The survival of a mutant gene and the maintenance of polymorphism in subdivided populations', *Am. Nat.* **108**, 20–28.

Pollak, E. (2000), 'Fixation probabilities when the population size undergoes cyclical fluctuations', *Theor. Pop. Biol.* **57**, 51–58.

Pulliam, H. R. (1988), 'Sources, sinks, and population regulation', *Am. Nat.* **132**, 652–661.

Pulliam, H. R. (2000), 'On the relationship between niche and distribution', *Ecol. Lett.* **3**, 340–362.

Redner, S. (2001), *A Guide to First-Passage Processes*, 1st edn, Cambridge: Cambridge University Press.

Reimann, P., Schmid, G. J. & Hänggi, P. (1999), 'Universal equivalence of mean first-passage time and kramers rate', *Phys. Rev. E* **60**, R1.

Richter-Dyn, N. & Goel, N. S. (1972), 'On the extinction of a colonizing species', *Theor. Pop. Biol.* **3**, 406–433.

Risken, H. (1996), *The Fokker-Planck equation*, 3rd edn, Berlin: Springer-Verlag.

Ronce, O. & Kirkpatrick, M. (2001), 'When sources become sinks: migrational meltdown in heterogeneous habitats', *Evolution* **55**, 1520–1531.

Rosenzweig, M. L. (1987), 'Habitat selection as a source of biological diversity', *Evol. Ecol.* **1**, 315–330.

Roughgarden, J. (1971), 'Density-dependent selection', *Ecology* **52**, 453–468.

Roughgarden, J. (1979), *Theory of Population Genetics and Evolutionary Ecology: An Introduction*, 1st edn, New York: MacMillan.

Rouhani, S. & Barton, N. H. (1987), 'The probability of peak shifts in a founder population', *J. Theor. Biol.* **126**, 51–62.

Saccheri, I. & Hanski, I. (2006), 'Natural selection and population dynamics', *Trends Ecol. Evol.* **21**, 343–47.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W. & Hanski, I. (1998), 'Inbreeding and extinction in a butterfly metapopulation', *Nature* **392**, 491–494.

Scahffer, M. L. (1987), Minimum viable populations: coping with uncertainty, *in* M. E. Soulé, ed., 'Viable Populations for Conservation', New York: Cambridge University Press, pp. 69–86.

Shpak, M. & Kondrashov, A. S. (1999), 'Applicability of the hypergeometric phenotypic model to haploid and diploids populations', *Evolution* **53**, 600–604.

Singh, R. S. & Morton, R. A. (2004), Beyond beanbag genetics: Wright's adaptive landscape, gene interaction networks, and the evolution of new genetic systems, *in* R. S. Singh & M. K. Uyenoyama, eds, 'The Evolution of Population Biology', Cambridge: Cambirdge University Press.

Slatkin, M. (1973), 'Gene flow and selection in a cline', *Genetics* **75**, 733–756.

Slatkin, M. (1977), 'Gene flow and genetic drift in a species subject to frequent local extinctions', *Theor. Pop. Biol.* **12**, 253–262.

Slatkin, M. (1978), 'Spatial patterns in the distribution of polygenic characters', *J. Theor. Biol.* **70**, 213–228.

Slatkin, M. (1981), 'Fixation probabilities and fixation times in a subdivided population', *Evolution* **35**, 477–488.

Slatkin, M. (1985), 'Gene flow in natural populations', *Ann. Rev. Ecol. Syst.* **16**, 393–430.

Slatkin, M. (1987), 'Gene flow and geographic structure of natural populations', *Science* **236**, 787–792.

Slatkin, M. (1994), 'Linkage disequilibrium in growing and stable populations', *Genetics* **137**, 331–336.

Slatkin, M. & Maynard Smith, J. (1979), 'Models of coevolution', *Quart. Rev. Biol.* **54**, 233–263.

Soulé, M. E. & Simberloff, D. (1986), 'What do genetics and ecology tell us about the design of nature preserves?', *Biol. Conser.* **35**, 18–40.

Spichtig, M. & Kawecki, T. J. (2004), 'The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments', *Am. Nat.* **164**, 70–84.

Spielman, D., Brook, B. W. & Frankham, R. (2004), 'Most species are not driven to extinction before genetic factors impact them', *Proc. Natl. Acad. Sci.* **101**, 15261–15264.

Spieth, P. T. (1974), 'Gene flow and genetic differentiation', *Genetics* **78**, 961–965.

Stacey, P. B., Johnson, V. A. & Taper, M. L. (1997), Migration within metapopulations: the impact upon local population dynamics, *in* I. Hanksi & M. Gilpin, eds, 'Metapopulation Biology: Evolution, Ecology and Genetics', San Diego: Academic Press, pp. 267–291.

Stearns, S. C. & Sage, R. D. (1980), 'Maladaptation in a marginal population of the mosquito fish *Gamusia affnis.*', *Evolution* **34**, 65–75.

Stephens, P. A., Sutherland, W. J. & Freckleton, R. P. (1999), 'What is the allee effect?', *Oikos* **87**, 185–190.

Stockwell, C. A., Hendry, A. P. & Kinnison, M. T. (2003), 'Contemporary evolution meets conservation biology', *Trends Ecol. Evol.* **18**, 94–101.

Strogatz, S. H. (2004), *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry and engineering (studies in nonlinearity)*, Cambridge: Cambirdge University Press.

Talkner, P. (1987), 'Mean first passage time and lifetime of a metastable state', *Z. Phys. B* **68**, 201–207.

Thoday, J. M. & Thompson, J. N. J. (1976), 'The number of segregating genes implied by continuous variation', *Genetica* **46**, 335–344.

Thomas, C. D. & Kunin, W. E. (1999), 'The spatial structure of populations', *J. Anim. Ecol.* **68**, 647–657.

Thompson, J. M. T. & Stewart, H. B., eds (1986), *Nonlinear Dynamics and Chaos*, Chichester: John Wiley and Sons.

Thompson, J. N. (1998), 'Rapid evolution as an ecological process', *Trends Ecol. Evol.* **13**, 329–332.

Tier, C. & Hanson, F. B. (1981), 'Persistence in density dependent stochastic populations', *Math. Biosci.* **53**, 89–117.

Tufto, J. (2001), 'Effects of releasing maladapted individuals: a demographic-evolutionary model', *Am. Nat.* **158**, 331–340.

Turelli, M. & Barton, N. H. (1990), 'Dynamics of polygenic characters under selection', *Theor. Pop. Biol.* **38**, 1–57.

Turelli, M., Barton, N. H. & Coyne, J. A. (2001), 'Theory and speciation', *Trends Ecol. Evol.* **16**, 330–343.

Van Kampen, N. G. (2007), *Stochastic Processes in Physics and Chemistry*, 3rd edn, Amsterdam: Elsevier.

Van Tienderen, P. H. (1991), 'Evolution of generalists and specialists in a spatially heterogeneous environment', *Evolution* **45**, 1317–1331.

Via, S. & Lande, R. (1985), 'Genotype-environment interaction and the evolution of phenotypic plasticity', *Evolution* **39**, 505–522.

Via, S. & Lande, R. (1987), 'Evolution of genetic variability in a spatially heterogenous envirionment: Effects of genotype-environment interactions', *Genet. Res.* **49**, 147–156.

Vuilleumier, S., Yearsley, J. M. & Perrin, N. (2008), 'The fixation of locally beneficial alleles in a metapopulation', *Genetics* **178**, 467–475.

Wade, M. J. & McCauley, D. E. (1988), 'Extinction and recolonization: their effects on the genetic differentiation of local populations', *Evolution* **45**, 995–1005.

Wallace, B. (1975), 'Hard and soft selection revisited', *Evolution* **29**, 465–473.

Walsh, J. B. (1984), 'Hard lessons for soft selection', *Am. Nat.* **124**, 518–526.

Watkinson, A. R. & Sutherland, W. J. (1995), 'Sources, sinks and pseudo-sinks', *J. Anim. Ecol.* **64**, 126–130.

Weiss, G. H. (1986), 'Overview of theoretical models for reaction rates', *J. Stat. Phys.* **42**, 3–36.

Whitlock, M. C. (1997), 'Founder effects and peak shifts without genetic drift: adaptive peak shifts occur easily when environments fluctuate slightly', *Evolution* **51**, 1044–1048.

Whitlock, M. C. (2003), 'Fixation probability and time in subdivided populations', *Genetics* **164**, 767–779.

Whitlock, M. C. & Barton, N. H. (1997), 'The effective size of a subdivided population', *Genetics* **146**, 427–441.

Whitlock, M. C. & McCauley, D. E. (1990), 'Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups', *Evolution* **816**, 97–169.

Whitlock, M. C. & McCauley, D. E. (1999), 'Indirect mreasures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$', *Heredity* **82**, 117–125.

Wiens, J. J. & Graham, C. H. (2005), 'Niche conservatism: integrating evolution, ecology and conservation biology', *Annu. Rev. Ecol. Evol. Syst.* **36**, 519–539.

Willi, Y., van Bushrick, J. & Hoffman, A. A. (2006), 'Limits to the adaptive potential of small populations', *Annu. Rev. Ecol. Evol. Syst.* **37**, 433–458.

Wolfram, S. (1999), *Mathematica*, ver. 4.0 edn, Cambridge: Cambirdge University Press.

Wright, S. (1931), 'Evolution in Mendelian populations', *Genetics* **16**(2), 97–159.

Wright, S. (1932), 'The roles of mutation, inbreeding, crossbreeding and selection in evolution', *Proc. 6th Int. Cong. Genet.* **1**, 356–366.

Wright, S. (1940), The statistical consequences of mendelian heredity in relation to speciation, *in* J. S. Huxley, ed., 'The New Systematics', Oxford: Oxford University Press, pp. 161–183.

Wright, S. (1968-78), *Evolution and the Genetics of Populations*, Vol. 1-4, 1st edn, Chicago: Chicago University Press.

Wright, S. (1986), *Evolution: Selected Papers*, Chicago: Chicago University Press.